DOCUMENT RESUME

ED 087 820 TM 003 453

| | |
|---|---|
| AUTHOR | Freeberg, Norman E. |
| TITLE | Development of Assessment Measures for Use With Youth-Work Training Program Enrollees Phase II: Longitudinal Validation. |
| INSTITUTION | Educational Testing Service, Princeton, N.J. |
| SPONS AGENCY | Manpower Administration (DOL), Washington, D.C. Office of Research and Development. |
| REPORT NO | DLMA-41-9-005-32-1; ETS-PR-74-1 |
| PUB DATE | Jan 74 |
| NOTE | 73p. |
| EDRS PRICE | MF-$0.65 HC-$3.29 |
| DESCRIPTORS | Adolescents; Aspiration; Attitude Tests; *Disadvantaged Groups; Employment Opportunities; Evaluation Techniques; Interest Tests; Job Satisfaction; Job Skills; Longitudinal Studies; Motivation; *Out of School Youth; Personality Tests; Social Adjustment; Test Reliability; *Tests; *Test Validity; Vocational Adjustment; Work Attitudes; *Work Experience Programs |
| IDENTIFIERS | *Enrollee Test Battery; Neighborhood Youth Corps |

ABSTRACT

A battery of tests, designed expressly for use in assessment of disadvantaged adolescents enrolled in youth-work training programs, was validated predictively against both short-term (program completion) and longer-term (post-program) behavioral outcome dimensions. Measures of the battery displayed their best validities with such longer-term criterion dimensions as Social and Vocational Adjustment, Job Success and Satisfaction and Job Search Motivation. The most valid tests were three Practical Reasoning measures (i.e., skill at following directions in job tasks); three Vocational Orientation measures (i.e., knowledge of job requirements and skill at seeking and holding jobs) and two attitudinal measures (Self-Esteem and Attitude Toward Authority). The most relevant short-term criteria were also identified and future potential of the test battery for programmatic evaluation and guidance applications was discussed. (Author)

ED 087820

PR-74-1

# DEVELOPMENT OF ASSESSMENT MEASURES FOR USE WITH YOUTH-WORK TRAINING PROGRAM ENROLLEES PHASE II: LONGITUDINAL VALIDATION

Norman E. Freeberg

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

January 1974

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

ED 087820

ETS PR-74-1

# DEVELOPMENT OF ASSESSMENT MEASURES FOR USE WITH YOUTH-WORK TRAINING PROGRAM ENROLLEES

## PHASE II: LONGITUDINAL VALIDATION

Norman E. Freeberg

Educational Testing Service

Princeton, N. J.

Final Report
U.S. Dept. of Labor Contract #41-9-005-32
January, 1974

This report was prepared for the Manpower Administration, U.S. Department of Labor, under research and development contract No. 41-9-005-32. Since contractors conducting research and development projects under Government sponsorship are encouraged to express their own judgment freely, this report does not necessarily represent the official opinion or policy of the Department of Labor. The contractor is solely responsible for the contents of this report.

| BIBLIOGRAPHIC DATA SHEET | 1. Report No. DLMA 41-9-005-32-1 | 2. | 3. Recipient's Accession No. |
|---|---|---|---|

4. Title and Subtitle

DEVELOPMENT OF ASSESSMENT MEASURES FOR USE WITH YOUTH-WORK TRAINING PROGRAM ENROLLEES. PHASE II: LONGITUDINAL VALIDATION

5. Report Date: January, 1974

6.

7. Author(s): Norman E. Freeberg

8. Performing Organization Rept. No.

9. Performing Organization Name and Address

Educational Testing Service
Rosedale Road
Princeton, New Jersey 08540

10. Project/Task/Work Unit No.

11. Contract/Grant No.: DL 41-9-005-32

12. Sponsoring Organization Name and Address

U.S. Department of Labor
Manpower Administration
Office of Research and Development
1111 20th St., N.W., Washington, D.C. 20210

13. Type of Report & Period Covered: Final

14.

15. Supplementary Notes

16. Abstracts

A battery of tests, designed expressly for use in assessment of disadvantaged adolescents enrolled in youth-work training programs, was validated predictively against both short-term (program completion) and longer-term (post-program) behavioral outcome dimensions. Measures of the battery displayed their best validities with such longer-term criterion dimensions as Social and Vocational Adjustment, Job Success and Satisfaction and Job Search Motivation. The most valid tests were three Practical Reasoning measures (i.e., skill at following directions in job tasks); three Vocational Orientation measures (i.e., knowledge of job requirements and skill at seeking and holding jobs) and two attitudinal measures (Self-Esteem and Attitude Toward Authority).

The most relevant short-term criteria were also identified and future potential of the test battery for programmatic evaluation and guidance applications was discussed.

17. Key Words and Document Analysis. 17a. Descriptors

Psychological Tests

17b. Identifiers/Open-Ended Terms

Youth-work training programs

17c. COSATI Field/Group 5I

| 18. Availability Statement | 19. Security Class (This Report) | 21. No. of Pages |
|---|---|---|
| Distribution is unlimited. Available from National Technical Information Service, Springfield, Va. 22151. | UNCLASSIFIED | |
| | 20. Security Class (This Page) UNCLASSIFIED | 22. Price |

## Table of Contents


| | Page |
|---|---|
| Introduction | 1 |
| Method | 4 |
| Description of the Measures | 4 |
| Sample and Data Collection | 9 |
| Sampling Influences | 13 |
| Data Analysis | 14 |
| Results | 17 |
| Predictive Validity of the Enrollee Test Battery | 17 |
| Criterion Relevance | 38 |
| Effects of Sample Attrition | 43 |
| Conclusions | 46 |
| Implications for Future Use of the Measures | 55 |
| Appendix A. Program Completion Criterion Variables | 63 |
| Appendix B. Post-Program Criterion Variables | 65 |
| Appendix C. Formula for Standardized Factor Scores | 68 |

## Acknowledgements

The author extends his appreciation to the many Neighborhood Youth Corps project directors and other project staff who allowed the use of their time and their facilities. By general consensus, anonymity of those participating projects has been preserved.

Also acknowledged are several Educational Testing Service staff members who were helpful in various phases of the study: Mr. Samuel Barnett who served as data collection coordinator; Dr. Donald Rock for advice in data analyses and Ms. Inge Stiebritz for data organization and computer programming.

# Development of Assessment Measures for Use with Youth-Work Training Program Enrollees

## Phase II: Longitudinal Validation


## Abstract

A battery of tests, that had been designed for use with disadvantaged adolescents enrolled in youth-work training programs and had undergone preliminary development, was subjected to longitudinal, predictive validation. The tests, dealing with a variety of reasoning, attitudinal and vocational orientation constructs, were administered to Neighborhood Youth Corps enrollees who were then followed up at program completion and six months after completion, using previously developed criterion instruments.

Test validities were based on relationships between test scores and criterion factor scores obtained for short-term outcome measures at program completion and longer-term (post-program) performance outcomes. Scattered significant validities of low magnitude were found, using the short-term criteria, with knowledge and reasoning (cognitive) measures tending to have the more consistent patterns of validities--although against differing performance criteria for males and females.

It was with the longer-term outcomes that the best of the predictive validities could be found in test relationships to a variety of job adjustment, social adjustment and motivational dimensions. Tests measuring enrollee assessment of self-worth (Self-Esteem), knowledge of proper interpersonal behaviors on a job (Job Holding Skills), ability to follow directions in job-related reasoning tasks (Practical Reasoning), abilities involved in seeking employment (Job Seeking Skills) and a variety of Vocational Interest scales were among the most valid (r's in the mid .20's to high .30's).

Separate validity analyses for ex-enrollees who had never obtained full-time employment (and could thus be measured only with a limited range of non-job-related outcomes) also indicated the general superiority of cognitively-oriented predictors (i.e., Practical Reasoning and Job Seeking Skills) but, in this instance, for a criterion dimension of "Social Adjustment." Similarly, cognitive skill measures show some of the best validities when using a dichotomous criterion of post-program employment (i.e., No Full-Time Employment vs. Full-Time Employment).

In an examination of criterion equivalence, 13 of the most relevant outcome measures, usable at program completion, were identified. By virtue of their number and magnitude of significant correlations with longer-term, post-program criteria, such short-term measures as counselor ratings, work supervisor ratings, peer ratings, enrollee feelings of vocational adequacy and amount of trouble with police, qualify as a most defensible set of program completion outcomes, or dependent variables, for any future research, evaluation or test validation uses.

# DEVELOPMENT OF ASSESSMENT MEASURES FOR USE WITH YOUTH-WORK TRAINING PROGRAM ENROLLEES

## PHASE II: LONGITUDINAL VALIDATION

## INTRODUCTION

In an initial phase of development, a battery of measures was designed for disadvantaged adolescents of low verbal skill levels enrolled in youth-work training programs (Freeberg, 1968, 1970). Those instruments utilized an approach to content, format and method of administration that differs somewhat from conventional paper and pencil tests. They also covered a wide range of attitudinal and cognitive constructs considered pertinent to curricula of work-training programs (e.g., vocational orientation in the form of job knowledge, interests, plans, motivations, and aspirations; skills required to seek and maintain employment; personal and social perceptions involving self-esteem, willingness to defer gratification and attitudes toward authority figures).

Thirteen test booklets which make up the battery (designated as the "Enrollee Test Battery" or ETB) had been administered to 256 male and female high school dropouts, from poverty-level backgrounds, enrolled in the Neighborhood Youth Corps (N.Y.C.) Out-of-School Program (Levitan & Mangum, 1969). Measurement properties of the battery were examined: (a) at the item level (e.g., item difficulties, item-total scale correlations), (b) at the level of total scale scores, in terms of reliabilities (as internal consistency) and concurrent validities (using guidance counselor and work-site supervisor ratings as criteria), and (c) at the level of the total battery, on the basis of underlying dimensions or factor patterns. Those

analyses revealed test characteristics sufficiently acceptable to warrant the continuing efforts necessary to derive a professionally defensible set of measures, particularly the requirement for broader concurrent and predictive validation. Accordingly, additional concurrent validity was sought in a study of the relationships between measures of the ETB and a set of more conventional, formally published, paper and pencil tests (Freeberg & Reilly, 1972). Six measures, chosen as concurrent criteria, were essentially of the cognitive variety and were administered along with the ETB to a sample of 76 high school dropouts enrolled in N.Y.C. Those conventional measures ran the gamut from a paragraph comprehension test to vocabulary, arithmetic, figure analogies, maze learning, and rote memory skills. The correlations obtained reflected a substantial degree of concurrent validity for the cognitive skill measures of the ETB and the relative independence of the attitude and interest scales.

Beyond the demonstration of concurrent validity, it had always been anticipated that the predictive validity of the ETB must eventually undergo examination--a task that first required the ability to choose and apply multiple performance criteria, or measures of behavioral outcomes, most relevant to youth-work training program goals. Unfortunately, information regarding useful criterion measures was sparse. In essence, no studies were known of in which there were systematic attempts at empirical development of criteria for youth-work training programs (or for any class of manpower training programs), despite previous prodding and a clear rationale for their need (Borus & Tash, 1970).

As a first step toward identifying a core of relevant outcome measures that might serve test validation purposes--or a variety of research and

evaluation needs--a criterion development study was undertaken. Multiple outcome measures applicable at different time periods (i.e., short-term program completion outcomes and longer-term post-program outcomes) were first rationally defined, based on such sources as enabling legislation and interviews with training program professionals, and then incorporated in questionnaires designed for oral administration to enrollees (and former enrollees) on an individual basis. Cross-sectional samples, composed of N.Y.C. enrollees who were completing the training program and former enrollees who had been out of the program for a period of six months to about 1 1/2 years, were chosen from six cities throughout the country. Major dimensions of performance objectives were identified in terms of independent, coherent clusters of variables that could be used to define short-term "success" (e.g., dimensions of training program adjustment, social-community adjustment, job planning competency) and longer-term objectives (e.g., dimensions of overall job success and adjustment, aspiration for vocational success, community and family adjustment). Based on the results, modifications were made in data gathering procedures, questionnaire content and scoring, so that the most satisfactory of the criterion methods and materials could then be applied.

Purpose of the Study

Given the Enrollee Test Battery with some evidence of suitable measurement properties along with coherent criterion variables having known dimensional attributes, the next logical research step is to incorporate these tests and criteria in a longitudinal validation study. Determination of the levels of predictive validity for each measure of the ETB against the program completion and post-program behavioral outcomes, constitutes the primary purpose of the present study. Such information would serve as the basis for

identifying those tests of the ETB most acceptable for any desired assessment purposes in the youth-work training setting.

A secondary, but important, study purpose realizable with a longitudinal design has to do with continued development of criteria and is aimed at determining the relevance of short-term, program-completion criteria based on their relationships to the longer-term post-program criterion measures. The often cited advantages of being able to identify the most relevant short-term program outcomes reside in the capability to substitute less costly and more readily available variables for the more difficult to obtain, longer-term (post-program) forms of outcome information (Thorndike, 1949; Wherry, 1957). Evidence that can be used to select the most relevant of short-term behavioral criterion measures can thus facilitate and improve the conduct of whatever research, evaluation or test validation procedure may be undertaken with training program enrollees.

As an adjunct to a longitudinal study of multiple predictors (tests) and multiple criteria, there is an additional opportunity to describe meaningful aspects of enrollee characteristics and performance that follow training program participation (i.e., what happened to former enrollees in terms of areas of later activities and achievements?). Similarly those criterion dimensions and their patterns of predictability can provide clues to particular areas of enrollee performance for which the training programs should strive to achieve more effective change.

## METHOD

### Description of the Measures

The predictor measures of the Enrollee Test Battery are described briefly below, followed by a summary discussion of the short-term (program completion)

and long-term (post-program) categories and variables of the criterion questionnaires. Although detailed knowledge of the rationale, design and content of those instruments requires some familiarity with prior studies (Freeberg, 1968, 1970; Freeberg & Reilly, 1971), these concise descriptions should be appropriate for a general understanding of the method and results to be presented here.

A. The Enrollee Test Battery (ETB) is made up of 13 test booklets for males that yield 21 scores and 12 booklets for females that yield 20 scores. These relatively short (10 to 28 item) test booklets contain both verbal and pictorial information, and are intended for oral administration to small groups of 12 or fewer respondents. Of the booklets, three are essentially cognitive in content and intended to measure practical reasoning skill in job-related tasks:

(1) Practical Reasoning: Zip Coding - PRZ (10 items): provides the respondent with information needed for a hypothetical post office job requiring him to sort mail using zip codes. Questions are presented in a multiple-choice format pertaining to that task.

(2) Practical Reasoning: Map Reading - PRM (10 items): presents a map illustrating several square city blocks, along with information needed to answer multiple choice questions about delivering materials to different locations shown on the map.

(3) Practical Reasoning: File Card Sorting - PRF (10 items): provides the respondent with information printed on file cards to be used in a clerical task requiring selection of potential job applicants who possess specified characteristics and skills. Questions in multiple-choice format require selection of applicant file cards with appropriate information in appropriate order. (The measure is utilized with male

respondents only since, in previous analyses, the item difficulty levels were far too low for females.)

Four other measures, dealing with aspects of vocational orientation, consist of:

(4) Job Knowledge - JK (27 items): requires answers to multiple-choice items having to do with requirements for a variety of jobs (e.g., carpenter, auto mechanic, policeman) in terms of their educational requirements, starting salary, primary task performed, hours of work, place where work is performed and tools utilized.

(5) Job Seeking Skills - JSS (17 items): presents multiple-choice items dealing with ways of looking for jobs, some of which entail interpretation of newspaper want ads and portions of job application blanks.

(6) Job Holding Skills - JHS (11 items): depicts situations that might be encountered on a job with regard to supervisors' requests, appropriate dress, punctuality, etc., for which the respondent chooses the reply he would give, or the most appropriate behavior called for, in that situation.

(7) Level of Aspirations and Level of Plans - LOA; LOP (16 items per scale): are obtained from the same sets of items which pictorialize jobs and designate the job titles. The respondent indicates the degree to which he would aspire (prefer) to engage in each occupation and then the degree to which it is likely that he would enter that same occupation. The discrepancy between the score for aspirations and the score for plans (LOA - LOP) was also utilized as a separate score.

There are five measures, classified as attitudinal, which deal with the following constructs:

(8) Attitude toward Authority - AA (12 items): in which scenes depicting responses of an adolescent to authority figures (e.g., teacher, policeman, parents, judge, etc.) are used to elicit the degree of pro- or anti-authority feelings by the respondent.

(9) Self-Esteem - SE (16 items): depicts situations pictorially in which the respondent indicates the degree to which he feels himself "worthy" of the desirable outcomes (getting a diploma, job, promotion, etc.).

(10) Motivation for Vocational Achievement - MVA (10 items): presents statements (item stems) that bear on the desirability of seeking and maintaining employment. These require the individual to respond (Likert-Scale format) in terms of the degree of agreement with the statement.

(11) Deferred Gratification - DG (16 items): presents statements dealing with the willingness to delay present reward for future gain; toward which the respondent indicates his degree of agreement.

(12) Enrollee Rating Scale - ERS (15 items): uses verbal items in Likert-Scale format which deal with the degree of suitability of program professionals (i.e., counselors and work supervisors) and other aspects of the training program in general, as evaluated by the enrollee.

The remaining test booklet in the ETB provides scores on seven interest scales:

(13) Interest in Vocational Tasks - VI (28 items): presents pictorial and verbal information regarding job tasks, characteristic of a number of occupations. The respondent indicates the degree to which he would favor performing each task. Separate scale scores are derived for each of seven interest categories (4 items each) designated as Aesthetic, Clerical, Service, Technical, Outdoor, Science, and Business.

Included in all analyses along with the measures of the ETB is a grade level reading score obtained from either the California Achievement Test (Tiegs & Clark, 1970) or the Stanford Achievement Test (Kelley, Madden, Gardner, & Rudman, 1964). The score is intended to provide tentative comparisons between validities for a conventional index of verbal-intellectual skill and the subtests of the ETB. However, considerable caution will be required in interpreting results for this score, since there was no attempt made to equate the grade scores from each of the two reading tests and, more important, the information was available for only 160 enrollees of 563 originally tested with the ETB. The severely reduced sample sizes available for the score, over the course of the longitudinal follow-up, will be indicated appropriately in all tabled results.

B. Criterion Measures: Performance outcomes were incorporated in two separate questionnaires. One is designated as the "Program Completion Questionnaire" and deals with variables reflecting short-term measures of "success" appropriate to enrollee performance at about the time that the enrollee has completed his period of training program participation (Appendix A provides an explanatory listing of the 32 short-term criterion variables used). These variables are confined largely to such enrollee-provided information as social and family adjustments, vocational expectations and intentions, and financial competency; with variables from external (other-than-enrollee) sources, based on judgmental ratings of enrollee competencies by others (i.e., counselors, work site supervisors, and peers) as well as number of work site assignments and absences during the training period.

The longer term outcome measures, found in the "Post-Program Questionnaire," are listed in Appendix B and encompass variables that deal primarily

with job-related experiences such as the seeking and obtaining of employment (e.g., methods used, amount of search effort expended), job performance (e.g., job skill level, salary level, raises, employer rating), and job satisfaction. Other items deal with vocational expectations and motivations, educational attainments or plans, and various questions intended to measure family and social adjustments.

All of the above job-related criterion information can, of course, apply only to a sample of former enrollees who had obtained some type of full-time employment. For those not able to do so (and, as will be seen, that group constitutes a significant proportion of former enrollees), post-program criterion variables are necessarily limited to aspects of work motivation and planning, along with social and family adjustment. For purposes of analysis, these individuals represent a separate post-program criterion group with a different configuration of applicable performance outcomes and for whom the predictive validities of the ETB must be examined independently.

Program Completion and Post-Program Questionnaires were designed for individual oral administration and employ a verbal "style" that is essentially an informal one. Either of the questionnaires can be administered in 20 to 30 minutes.

Sample and Data Collection

The longitudinal sample utilized for the study consisted--at the time of initial testing with the ETB--of 267 males and 296 females enrolled in the Neighborhood Youth Corps, Out-of-School Program. Those 563 enrollees were in the early stages of their training (i.e., they had been enrolled for approximately two to four weeks) and were tested at N.Y.C. projects located in 14

cities throughout the country. Six cities were located in the Eastern U.S., 4 in the Midwest, and 4 in the Far West.

Of that initial sample, 399 (184 males and 215 females) had remained enrolled for a sufficient period of time to be classified as having "completed" the program (i.e., approximately 5 months or more) and would also volunteer to respond to the Program Completion Questionnaire.[1]

In the post-program follow-up phase, 219 enrollees who had responded to the questionnaire at program completion could be located approximately six months after they had left the training program, and were also willing to respond to the Post-Program Questionnaire. Of the 219 members of the Post-Program criterion sample, 121 (59 males and 62 females) had obtained full-time employment and are designated as the "Post-Program:Employed" sample. The remaining 98 (45 males and 53 females) had not obtained full-time employment at any time during the six-month period and comprise the "Post-Program: Not Employed" sample.[2] Between 80% and 85% of the samples obtained at each data collection phase were Black, with approximately 10% to 15% of the samples of Spanish-speaking ethnic backgrounds (Puerto Rican or Chicano) and the remainder classifiable as White. Mean age of the group of 399 at the time they completed the program was 17.9 years.

---

[1]The time of enrollment that constitutes completion is somewhat arbitrary and is based solely, for present study purposes, on legislative guidelines that suggest six months as the nominal training period for the N.Y.C. program.

[2]The determination of Full-Time Employment was based on the individual's having held a paid job for 20 hours or more over a period of one week or longer.

Table 1 shows the sample sizes, by sex, over the course of the longitudinal study and the number of cities represented at each data collection phase.

The extensive sample attrition found was not unexpected in a program where dropout rates are known to fluctuate widely and the trainees are drawn from a "hard-to-locate" population (U. S. Dept. of Labor, 1969). Attrition occurred not only because of inability to locate respondents and enlist their cooperation, but also because of loss of all subsequent data in a given city when an interviewer was forced to withdraw from the project and a suitable substitute could not be found (as occurred in three of the 14 cities).

Data collection was to be carried out by one interviewer in each of the 14 cities who was responsible for all three data-gathering phases over a time period of some 12 to 15 months. This entailed not only administration of the ETB and the two follow-up questionnaires, but the gathering of information from sources other than the enrollee (e.g., training program records, employer ratings, police contacts, etc.). The 14 interviewers participated in three-day training sessions to acquaint them with the study purposes, to explain procedures for conducting the follow-up and to rehearse the administration of the ETB subtests and the criterion questionnaires. All interviewers were residents of the city in which they were to gather the data and most were minority group members.

Payments to the enrollees were made by the interviewers and consisted of $3.00 for responding to the Program Completion Questionnaire and $5.00 for the Post-Program Questionnaire. Interviewers were paid a lump sum for the initial testing phase and on a per-interview basis for administration of the two criterion questionnaires.

Table 1

Composition of Longitudinal Sample

| | Data Collection Phases | | |
|---|---|---|---|
| 1. Initial Testing | 2. Program Completion Criteria | 3. Post-Program Criteria | |
| | | (Employed Full-Time) | (Never Employed Full-Time) |
| Males: N = 267 | 184 | 59 | 45 |
| Females: N = 298 | 215 | 62 | 53 |
| Total: N = 565 | N = 399 | N = 121 | N = 98 |
| | | Total: N = 219 | |
| Total # Cities 14 | 13 | 11 | |

Sampling Influences

Pervasive problems generic to any efforts at obtaining personal information from poverty-level, inner city, minority-group youngsters, can serve to influence the representativeness of the sample obtained. Limitations in control over sample selection introduce likely sources of bias that are worth noting. For the most part, samples obtained over the course of this longitudinal follow-up must be viewed as samples-of-opportunity, since participation on the part of any training program project (i.e., the project director and project professionals), or of individual enrollees and former enrollees, is entirely voluntary--as is their participation in the program itself--and needed cooperation is not uniformly elicited. In addition, selection of data-gathering locales (urban areas) is limited by the very availability of suitable professionals willing to expend the effort and assume the risks of tracing the whereabouts of adolescents or young adults residing in ghetto areas of large cities. By sufficient breadth of geographic areas sampled, use of monetary incentives and application of procedures culled from prior experience with "hard-to-locate" populations (U. S. Department of Labor, 1969), it is hoped that sampling biases can be reduced.

Experience from previous data collection efforts with tests and questionnaires, for example, have indicated a tendency for biases to be introduced in the form of disproportionally larger samples of females being obtained. The females are simply more willing to comply with requests for test participation and are more easily located in follow-up attempts. Anticipation of the problem by the planning of special efforts to locate male enrollees must, therefore, be made and in the case of the present study seem to have met with reasonable success. Nevertheless, the sheer magnitude of sample loss experienced, over

time, for poverty level, minority group, young-adult populations and the possible selective nature of that attrition, with regard to what may be unidentified but important behavioral characteristics, must remain a subject of concern until such effects can be thoroughly studied. Limited examination of variables that might bear on selective attrition is possible with the data available in the present study and will be considered briefly in the Results section.

Data Analysis

In order to demonstrate the level and pattern of validities for each subtest of the Enrollee Test Battery, it was first necessary to summarize the sets of numerous criterion variables in terms of broader constructs (i.e., factors) and from those, to derive composite scores. This was accomplished by obtaining separate intercorrelation matrices for the 32 program completion criterion variables, the 40 variables of the Post-Program:Employed sample, and the 29 variables used with the Post-Program:Not Employed sample. The three matrices were then factor analyzed using a principal components solution with varimax rotation to orthogonality (Kaiser, 1958). This was the same analytical procedure used, with almost identical sets of variables, to obtain criterion dimensions in the previously cited cross-sectional criterion development study (Freeberg & Reilly, 1971), which allows for general comparisons between those results and the present longitudinal study sample.

With the availability of a more manageable number of relatively independent, interpretable criterion constructs (factors), a single score could then be obtained for each factor. Essentially it is a standardized factor score, the description of which is shown in Appendix C. Finally, correlations (validity coefficients) were computed between each of the 21 ETB test scores (plus reading grade level) and the factor scores that had been obtained for each criterion dimension.

The analyses were performed separately, by sex, for the program completion (short-term) criterion data, because sample sizes (184 males and 215 females) were considered acceptable for factor analyses. Unfortunately, division of samples by sex and extraction of factors for either of the post-program samples would be highly questionable on the basis of available sample sizes in relation to the number of variables involved. Thus, longer-term test validities were computed only for the total groups (i.e., sexes combined).

In order to determine the relevance of the short-term criteria, the 32 program completion variables were correlated with each of the 40 long-term, post-program variables for the employed sample of former enrollees. The resulting 32 x 40 rectangular matrix was summarized by selecting short-term criteria with the largest number of significant r's and presenting them, in ranked order, along with the significant coefficients, as a way of highlighting the most predictive (relevant) of the short-term outcomes.

Possible sampling biases, introduced by attrition over the three data-gathering phases, were assessed by contrasting differences between mean scores on the Enrollee Test Battery for those members of the sample who were able to be followed up six months after completion of the program ("Responders") and those who could not be located or who refused to participate ("Non-Responders").

Missing data occurred throughout the samples, requiring missing-data analyses for virtually all computations. For example, there were considerable reductions in sample size for variables of police contacts, employer ratings, and peer ratings, because of constraints on information availability, in addition to relatively small, but widely scattered, enrollee refusals to respond to individual questionnaire items. A major reduction in the sample

of Reading Level scores resulted from resistance on the part of interviewers to administer what they considered undesirable ("culturally biased"?) reading tests to minority group, school dropouts and thus risk their credibility with the youngsters for the later follow-up efforts.

## RESULTS

The presentation of results will be in three subsections: The first is devoted to ETB predictive validities for the program completion and post-program criterion samples (including the factor analyses of criteria required as a step in that process). The second will deal with the extent of criterion relevance, or equivalence, between short- and long-term criteria and the third subsection will present the comparisons of ETB test scores for follow-up "responders" and "non-responders."

### Predictive Validity of the Enrollee Test Battery

A. Short-Term (Program Completion) Validity: The factor designations and loading patterns for the program completion criteria are presented in Table 2 for males and Table 3 for females and, for present study purposes, require only summary discussion. For each subsample, six factors had been extracted and rotated. In each case, four of the factors were found to be reasonably interpretable, with a sufficient number of loadings of sufficient magnitude (.30 or greater) to warrant inclusion here. These are shown in the order in which they were extracted. Thirty-four percent of the total variance of the intercorrelations was accounted for in the male sample by these four factors and 33% by the four factors reported for the female sample.

A factor of Training Program Adjustment (Factor I for females; Factor III for males) appears in both samples--although a more dominant factor for females--with the lead loadings in each case appearing for Counselor and Work Site Supervisor ratings. An outcome dimension clearly reflecting a form of Social Adjustment (Factor II for males; Factor III for females) is also found in both samples, with the similarity defined by the largest loadings

Table 2

Rotated Factor Loadings for Program Completion Criteria
(Males: N = 182)

| Criterion Variables | Factors | | | |
|---|---|---|---|---|
| | I Positive Vocational & Social Attitudes | II Social Adjustment | III Training Program Adjustment | IV Vocational Planning Competency |
| 1. G.E.D. | .32 | .00 | .41 | .15 |
| 2. Feelings of Voc. Adeq. | .64 | -.10 | .12 | .33 |
| 3. 1st Move | .05 | .10 | -.09 | .43 |
| 4. Reason 1st Move | -.11 | .47 | -.06 | .06 |
| 5. Awareness of Job Characteristics | .37 | .16 | .12 | .19 |
| 6. Job Knowledge (Self-Assessment) | .14 | .09 | -.32 | .17 |
| 7. Ability to Perform Job | .14 | -.07 | -.29 | .14 |
| 8. Starting Salary Expected | .14 | .00 | -.11 | .17 |
| 9. Top Salary Expected | .27 | .11 | -.25 | .13 |
| 10. Ways to Look for Job | -.35 | .12 | -.02 | .21 |
| 11. Imp. Ask Job Interviewer | .00 | -.07 | .22 | .55 |
| 12. Short Term Job Plans | -.09 | .03 | .01 | .46 |
| 13. Long Term Job Plans | .17 | .01 | .03 | .75 |
| 14. Ability to Accomp. Plans | .08 | .13 | -.01 | .66 |
| 15. Get Along with Family | .18 | .66 | .06 | -.07 |
| 16. Family Feel about Enrollee | .51 | .44 | .12 | -.04 |
| 17. Financial Assist. to Family | .14 | .25 | .00 | .09 |
| 18. Save Money | .59 | .12 | .05 | .01 |
| 19. # People Giving Hard Time | -.17 | -.60 | -.27 | -.11 |
| 20. Health Problems | .02 | .19 | .06 | .15 |
| 21. Trouble with Police | -.14 | -.52 | -.28 | .01 |
| 22. Imp. Keep out of Trouble | .55 | .39 | .07 | .06 |
| 23. # Jobs Chosen for Training | .09 | -.39 | .06 | .11 |
| 24. Willing Train Full Time | .37 | -.28 | .03 | -.09 |
| 25. Willing Train Part Time | .56 | -.11 | .21 | .01 |
| 26. Work Motivation | .34 | .08 | .31 | .08 |
| 27. # Work Sites | -.02 | -.03 | -.21 | .06 |
| 28. # Days Absent | .28 | -.03 | -.10 | -.01 |
| 29. Counselor Rating | .18 | .13 | .75 | .18 |
| 30. Work Supervisor Rating | .18 | .16 | .75 | .09 |
| 31. Peer Rating | .12 | -.04 | .55 | -.03 |
| 32. # Times Chos.for Peer Rat. | -.15 | .41 | -.08 | .14 |

Table 3

Rotated Factor Loadings for Program Completion Criteria
(Females: N = 215)

| Criterion Variables | Factors I Training Program Adjustment | II Work Motivation and Planning | III Personal-Social Adjustment | IV Positive Vocational Orientation |
|---|---|---|---|---|
| 1. G.E.D. | -.21 | .41 | -.00 | .34 |
| 2. Feelings of Voc. Adeq. | .30 | .19 | .41 | .21 |
| 3. 1st Move | .20 | .37 | .09 | .08 |
| 4. Reason 1st Move | -.03 | -.22 | -.00 | .52 |
| 5. Awareness of Job Characteristics | .16 | .12 | .24 | .43 |
| 6. Job Knowledge (Self-Assessment) | .20 | -.50 | .14 | .12 |
| 7. Ability to Perform Job | .06 | -.52 | .13 | .08 |
| 8. Starting Salary Expected | .09 | -.03 | -.11 | -.13 |
| 9. Top Salary Expected | -.03 | .02 | -.04 | -.04 |
| 10. Ways to Look for Job | -.01 | -.24 | -.03 | .40 |
| 11. Imp. Ask Job Interviewer | .37 | .04 | -.07 | .54 |
| 12. Short Term Job Plans | -.02 | .07 | .15 | .62 |
| 13. Long Term Job Plans | .38 | .52 | -.02 | .27 |
| 14. Ability to Accomp. Plans | .35 | .27 | .01 | .33 |
| 15. Get Along with Family | .03 | -.16 | .58 | .01 |
| 16. Family Feel about Enrollee | -.15 | .10 | .67 | .09 |
| 17. Financial Assist. to Family | .10 | -.06 | .14 | .21 |
| 18. Save Money | .03 | .33 | .13 | -.24 |
| 19. # People Giving Hard Time | -.32 | .24 | -.54 | .06 |
| 20. Health Problems | -.00 | .41 | -.19 | .17 |
| 21. Trouble with Police | .05 | .03 | -.12 | .07 |
| 22. Imp. Keep out of Trouble | -.07 | -.01 | .48 | .23 |
| 23. # Jobs Chosen for Training | -.05 | .43 | -.11 | .32 |
| 24. Willing Train Full Time | .16 | .13 | .00 | -.10 |
| 25. Willing Train Part Time | .09 | .52 | .18 | -.16 |
| 26. Work Motivation | -.04 | .65 | -.04 | -.01 |
| 27. # Work Sites | -.25 | .01 | -.15 | .09 |
| 28. # Days Absent | -.77 | -.02 | .09 | .02 |
| 29. Counselor Rating | .75 | -.08 | .16 | .06 |
| 30. Work Supervisor Rating | .76 | -.11 | .11 | .05 |
| 31. Peer Rating | .29 | -.08 | .52 | -.02 |
| 32. # Times Chos. for Peer Rating | .12 | .04 | .55 | -.27 |

in each sample on family adjustment variables, a community adjustment index (# people giving enrollee a hard time) and peer rating scores. For the two remaining factors of approximately similar designation there are probably as many differences as similarities in the item loadings. Thus, the dominant Positive Vocational and Social Attitudes factor for males is indicative of not only striving for improvement (e.g., working toward or having received the G.E.D.) and greater overall willingness to undergo job training, but also shades into areas of social adjustment (i.e., family adjustment, savings, importance of keeping out of trouble and feelings of vocational adequacy). For females the Positive Vocational Orientation factor is far more specific in focus, dealing primarily with job search and planning behaviors.

Planning ability, as found for males in Factor IV (Vocational Planning Competency) appears to be highly specific, covering almost exclusively an appreciation of what to anticipate or be aware of in looking for a job (i.e., loadings on ability to plan appropriately for jobs and what to ask about them in an interview). By contrast, among females, planning as an outcome dimension (Work Motivation and Planning) is more inextricably tied to motivation in the form of willingness to accept training, obtaining the G.E.D., and, predominantly, in expressions of work motivation (i.e., a lead loading on the "Work Motivation" Scale). Interestingly, it appears that females who show positive work motivation and planning behaviors tend to be those who are uncertain of knowledge about a desired job or their ability to perform it (loadings of -.50 and -.52 for variables #5 and #6). A similar pattern emerges for the loadings of those same two variables--although much weaker--with the male sample(as seen in Factor III).

Any direct comparisons of these longitudinal factorial results with the previously obtained program-completion dimensions of the cross-sectional study

(Freeberg & Reilly, 1971) are tenuous, since the previous analyses combined males and females in a single sample. However, it should be noted that general similarities in the major factors of Training Program Adjustment and Social Adjustment hold up despite the differences in study design and sample composition.

Of greater interest, as the primary focus of the present section, are the validities obtained for the ETB using factor scores from these dimensions as the criterion measures. Table 4 shows the short-term validities for males and Table 5 presents the validities for females. Ability to predict enrollee performance at the completion of the program with the ETB (predictive validity) is seen to be scattered over a variety of different types of measures, at relatively low levels of relationship, and to differ by sex. The best of the significant validities for males is with the Positive Vocational and Social Attitudes dimension for the ETB measures of "Self-Esteem" (r = .34) and "Level of Plans" (r = .23). Other noncognitive, attitudinal measures are conspicuous by their absence.

Some type of discernible pattern, for males, begins to emerge in the validities found with the Social Adjustment criterion dimension (Factor II). Here it would seem evident from the low but significant correlations that tests having known cognitive skill components (i.e., "Job Seeking Skills," "Practical Reasoning:Zip Coding," and "Practical Reasoning:File Card Sorting") tend to dominate, while only one noncognitive measure--"Job Holding Skills" (r = .21)--appears to have any value as a predictor of short-term social adjustment. No significant validities are found with the two remaining and less dominant criterion dimensions of the male sample.

The picture is not markedly improved by the validities obtained for the female sample (Table 5). The major definable pattern of significant, but low

Table 4

Validity Coefficients for Enrollee Test Battery with Program Completion Criterion Factor Scores

(Males: N = 182)

| Measures of Enrollee Test Battery | Criterion Factors | | | |
|---|---|---|---|---|
| | I Positive Voc. and Social Attitudes | II Social Adjustment | III Training Program Adjustment | IV Vocational Planning Competency |
| 1. Job Knowledge | .04 | .14 | -.07 | .04 |
| 2. Self-Esteem | .34* | .07 | .09 | .02 |
| 3. Deferred Gratification | .06 | -.06 | .10 | .14 |
| 4. Level of Aspiration | .11 | .09 | .06 | -.07 |
| 5. Level of Plans | .23* | .01 | .02 | -.14 |
| 6. Aspirations-Plans Diff. | -.14 | .11 | .03 | .00 |
| 7. Attitude Toward Authority | .00 | .09 | .14 | .03 |
| 8. Job-Seeking Skills | -.11 | .21* | -.01 | -.05 |
| 9. Job-Holding Skills | .11 | .21* | -.07 | .02 |
| 10. Motivation for Vocational Ach. | .15 | .13 | -.05 | .08 |
| 11. Practical Reasoning (PR):Map Read. | -.08 | .09 | -.09 | .01 |
| 12. PR: Zip Coding | -.01 | .20* | -.02 | .08 |
| 13. PR: File Card Sorting | -.02 | .19* | .03 | .07 |
| 14. Enrollee Rating Scale | .10 | .01 | .00 | .04 |
| 15. Reading Level (SAT, CAT)[a] | -.03 | .17 | .13 | -.21 |
| Vocational Task Interest | | | | |
| Aesthetic | .16 | -.04 | .04 | -.11 |
| Science | .09 | -.05 | .06 | -.06 |
| Technical | .07 | -.07 | .01 | -.09 |
| Clerical | .12 | -.09 | -.03 | -.05 |
| Business | .15 | -.09 | .02 | -.13 |
| Service | .14 | -.08 | .00 | -.03 |
| Outdoor | .01 | .06 | -.02 | -.11 |

[a] N = 79.

*r significant at .05 level.

Table 5

Validity Coefficients for Enrollee Test Battery with Program Completion Criterion Factor Scores

(Females: N = 215)

| Measures of Enrollee Test Battery | Criterion Factors I Training Program Adjustment | II Work Motivation & Planning | III Personal-Social Adjustment | IV Positive Vocational Orientation |
|---|---|---|---|---|
| 1. Job Knowledge | .17* | -.02 | -.10 | .15* |
| 2. Self-Esteem | -.03 | .06 | .01 | .07 |
| 3. Deferred Gratification | .08 | .13 | .11 | .03 |
| 4. Level of Aspiration | .01 | .11 | .10 | .03 |
| 5. Level of Plans | -.08 | .08 | .12 | .04 |
| 6. Aspirations-Plans Diff. | .07 | -.02 | .01 | -.07 |
| 7. Attitude Toward Authority | .18* | -.06 | -.06 | .00 |
| 8. Job-Seeking Skills | .18* | -.01 | -.06 | .19* |
| 9. Job-Holding Skills | -.02 | .00 | .08 | .16* |
| 10. Motivation for Vocational Ach. | -.05 | .01 | .03 | .01 |
| 11. Practical Reasoning (PR): Map Read. | .23* | -.06 | -.06 | -.02 |
| 12. PR: Zip Coding | .19* | -.09 | -.15* | .14 |
| 13. Enrollee Rating Scale | -.04 | -.06 | .03 | -.02 |
| 14. Reading Level (SAT, CAT)[a] | .24* | .08 | .10 | .40* |
| Vocational Task Interest | | | | |
| Aesthetic | -.06 | .03 | -.04 | -.07 |
| Science | .08 | .06 | .09 | .03 |
| Technical | -.01 | .04 | .14 | -.10 |
| Clerical | .11 | -.05 | .13 | .05 |
| Business | -.17* | .04 | .12 | -.01 |
| Service | -.01 | .01 | .17* | .00 |
| Outdoor | .03 | -.04 | .06 | -.08 |

[a] N = 64.

*r significant at .05 level.

levels of validity is found with the criterion dimension of Training Program Adjustment. There the two vocational orientation measures of "Job Knowledge" (r = .17) and "Job Seeking Skills" (r = .18) combine with the two Practical Reasoning measures ("Map Reading," r = .23; and "Zip Coding," r = .19) and Reading Level (r = .24) to give a clear edge to a variety of cognitive skills among females as predictors of this dominant dimension of Success at Program Completion. "Attitude Toward Authority" figures (r = .18) constitutes the single attitudinal measure significantly related to that adjustment criterion. The only other pattern of validities worth citing is a hint of some value for a group of ETB job-orientation tests ("Job Knowledge," "Job Seeking Skills," "Job Holding Skills") with the short-term outcomes that define a Positive Vocational Orientation dimension. Support for a suspicion that a degree of cognitive capability underlies the pattern is found in the respectable r of .40 between Reading Level and the criterion score.

In surveying the above validity results based on the short-term program-completion criteria, some question should be raised regarding the reasons for the generally poor showing of the ETB measures. Granted probable deficiencies in the quality of the criterion data (e.g., sampling and response inaccuracies), another form of explanation is that the period of five to six months following testing represents too short a time for a range of such complex behavioral constructs to predict a rather limited set of criterion measures that deal with enrollee self-report and assessments by others within training program confines. The more meaningful behavioral outcomes or accomplishments predictable by such measures may not manifest themselves until the enrollee is on his own and faced with the greater demands made in "real world" on-the-job performance. In part, some support for this contention will be evident in the later discussion of

results for the post-program validation. But, there is one other explanatory approach worth pursuing, which assumes that two distinct categories of adolescent enrollee enter the work-training program. One group is severely hampered by exceptionally poor verbal skills, at or near levels of functional illiteracy (probably accompanied by greater social maladjustments) and is unlikely to be changed significantly enough, by the short period spent in the program, to obtain subsequent employment. The other, higher level, group of enrollees is more likely to profit from the program or, perhaps, would obtain later employment in any event. Combining two groups of this sort with distinctly different capabilities may obscure the predictive value of the ETB measures for either group separately. In short, the assumption being made is that later program employment serves as a moderator variable (Ghiselli, 1963). This assumption is at least partially testable with available data and can be illustrated briefly using two of the more important program-completion criteria (counselor and work supervisor ratings) with the male sample. In Table 6 the validities of the ETB tests are shown with each of those two rating scales; first for the total male sample at program completion and then for the subsample of males who were later able to obtain full-time employment. For the total sample, there are a number of significant--if rather minimal--levels of validity, with a distinct superiority for the job orientation and cognitive skill types of measures. For the subgroup who are later employed, the number and magnitude of significant r's is seen to rise considerably. In effect they represent a more predictable group using a number of these ETB measures.[1]

---

[1]Without belaboring the point or taking up space for an additional table, there are, as might be expected, either no change or reductions in validities found when computed for a subsample of males who fail to obtain later employment. Furthermore, the general pattern of improved ETB validities appears to hold up (although not quite as dramatically) for a variety of the other criterion variables in addition to counselor and work supervisor ratings.

Table 6

Short-Term (Program Completion) Test Validities: Comparison of Total Male Sample with Subsample Later Employed

(Counselor and Work Supervisor Ratings as Criteria)

| Test | Total Program Completion Sample (Males: N = 155) | | Subsample Who Obtained Later Employment (Males: N = 54) | |
|---|---|---|---|---|
| | Criteria | | Criteria | |
| | Counselor Ratings | Work Sup. Ratings | Counselor Ratings | Work Sup. Ratings |
| 1. Job Knowledge | .10 | .16* | .23 | .33* |
| 2. Self-Esteem | .20* | .13 | .50* | .47* |
| 3. Deferred Gratification | -.01 | -.03 | .24 | .31* |
| 4. Level of Aspiration | .02 | .08 | -.02 | .08 |
| 5. Level of Plans | .07 | .03 | -.01 | .09 |
| 6. Aspirations-Plans Diff. | -.10 | .03 | -.05 | -.02 |
| 7. Attitude Toward Authority | .06 | .08 | .36* | .32* |
| 8. Job-Seeking Skills | .23* | .17* | .34* | .39* |
| 9. Job-Holding Skills | .20* | .24* | .36* | .54* |
| 10. Motivation for Vocational Ach. | .12 | .09 | .14 | .17 |
| 11. Prac. Reason (PR): Map Read. | .11 | .11 | .27* | .31* |
| 12. PR: Zip Coding | .19* | .18* | .34* | .33* |
| 13. PR: File Card Sorting | .21* | .15 | .28* | .30* |
| 14. Enrollee Rating Scale | -.05 | -.02 | -.04 | .05 |
| 15. Reading Level (SAT, CAT)[a] | .23* | .19 | .29 | .37* |
| Vocational Task Interest | | | | |
| 16. Aesthetic | .03 | .09 | .18 | .23 |
| 17. Science | .01 | .02 | .00 | .11 |
| 18. Technical | -.04 | .06 | -.04 | .08 |
| 19. Clerical | -.06 | .00 | -.10 | .04 |
| 20. Business | -.03 | -.03 | -.07 | .01 |
| 21. Service | -.07 | .00 | -.07 | -.01 |
| 22. Outdoor | .06 | .19 | .09 | .23 |

[a] N = 75 (Total Prog.Comp. Sample); N = 31 (Employed Subsample).

*r significant at .05 level.

Such a moderator of later employment certainly represents a post hoc, after the fact, influence that is of little practical value for present predictive validation purposes--unless there were some highly effective means for categorizing trainees at program entry into those who are, and are not, likely to achieve later employment, so that the measures of the test battery could be applied differentially. It does, however, represent a finding with broad implications for further research in attempting to define such differentiating enrollee characteristics and eventually for approaches to training program design. (The extent to which the present test battery is helpful in this regard is looked at briefly below, using later employment as a criterion variable.)

B. Longer-Term (Post-Program) Validity: As in the previous analyses, criterion factors were first extracted from the intercorrelation matrix of criterion variables and factor scores were then computed for each interpretable factor in order to provide a best summary criterion measure.

Intercorrelation of 40 variables, applied to the sample of former enrollees who obtained full-time employment (males and females combined), resulted in a matrix from which six factors were extracted and rotated. Five of those factors, accounting for 41% of the total variance, were found sufficiently interpretable to be included in the validity analyses. The five, as presented in Table 7, show an overwhelming vocational or job-related focus and again warrant only summary discussion. A notable difference, in contrasting these results with the factor patterns obtained from the cross-sectional study, is the failure of an independent social-personal adjustment dimension to appear. Instead the social aspect of longer-term adjustment is seen closely tied to a number of on-the-job vocational adjustment outcomes,

Table 7

Rotated Factor Loadings for Post-Program Criteria:
Employment Experience Group

(Total Sample: N = 122)

| Criterion Variables | I Overall Social and Vocational Adjustment (White Collar) | II On-the-Job Success and Satisfaction (Blue Collar) | III Job Search Motivation | IV Job Planning Competency | V Short-Range Job Orientation |
|---|---|---|---|---|---|
| 1. # Interviewer Contacts Required | -.05 | .12 | .08 | -.00 | .49 |
| 2. # Places Lived Since N.Y.C. | -.30 | .16 | .24 | .07 | .15 |
| 3. Employed Now (No/Yes) | .49 | .36 | -.08 | -.20 | -.08 |
| 4. Industrial Category-Wh/Bl.Collar | -.39 | .35 | -.33 | -.16 | -.05 |
| 5. Job Quality-Present or Last Job | .05 | .08 | -.17 | .48 | .02 |
| 6. Hours Worked/Week | -.14 | .52 | .22 | -.07 | .31 |
| 7. Time on Present (or Last) Job | .12 | .08 | -.33 | .10 | .44 |
| 8. # Jobs Held Since N.Y.C. | -.24 | .01 | .17 | -.24 | -.15 |
| 9. Time to Find 1st Job | .13 | -.08 | .27 | -.09 | -.59 |
| 10. # Sources to Find 1st Job | -.17 | .10 | .64 | .08 | -.25 |
| 11. # Places Interviewed | -.14 | -.04 | .83 | .06 | -.02 |
| 12. # Applications Filed | -.00 | .05 | .82 | .05 | .01 |
| 13. Met Job Expectations | .38 | .45 | -.33 | .20 | -.01 |
| 14. Starting Salary | -.00 | .78 | .06 | -.08 | .14 |
| 15. Salary Raise (No/Yes) | .13 | .26 | -.10 | -.06 | .34 |
| 16. Job Satisfaction | .42 | .51 | -.18 | .12 | -.00 |
| 17. Top Salary Expect. | -.21 | .67 | -.24 | .12 | .03 |
| 18. Level of Short-Term Plans | .07 | -.23 | .02 | .61 | -.17 |
| 19. Job Level Sought (Short Term) | .06 | -.10 | .10 | .81 | .04 |
| 20. Reason for Job Selection | .12 | .27 | .14 | .39 | .17 |
| 21. Knowledge of Job Sought (Self-Estimate) | .23 | .06 | -.09 | .05 | .66 |
| 22. Steps to Obtain Desired Job | .14 | .06 | .09 | .59 | .15 |
| 23. Level of Long Range Work Plans | -.01 | -.12 | -.01 | .57 | -.30 |
| 24. Amount School Attendance | .15 | -.36 | -.32 | .34 | -.09 |
| 25. Obtained G.E.D. Since N.Y.C. | .08 | .14 | -.06 | .22 | -.41 |
| 26. Getting Along with Family | .53 | .14 | -.12 | .08 | -.18 |
| 27. Family Feel About Enrollee | .62 | .28 | -.20 | .17 | -.27 |
| 28. Financial Assist. to Family | .47 | -.11 | -.02 | .10 | .08 |
| 29. # People Giving Hard Time | -.61 | -.16 | .05 | -.07 | -.05 |
| 30. Health Problems | .10 | -.04 | -.05 | -.18 | .10 |
| 31. # Times Visited Doctor | .04 | .06 | .10 | .02 | .25 |
| 32. Save Money | .43 | .17 | -.22 | .11 | -.38 |
| 33. Amount Trouble with Police | -.74 | .13 | .05 | -.07 | .02 |
| 34. Imp. Keep Out of Trouble | .66 | -.05 | .09 | -.07 | .16 |
| 35. Credit Buying | .36 | .34 | .08 | .07 | .11 |
| 36. Ease of Obtaining Credit | .49 | .16 | -.02 | .29 | -.16 |
| 37. # Visits to U.S.E.S. | -.01 | -.08 | .47 | -.07 | -.07 |
| 38. Permission for Employer Rating | .08 | .45 | .04 | -.18 | -.13 |
| 39. # Actual Police Contacts | -.73 | .20 | .02 | .01 | .03 |
| 40. Employer Rating | .73 | .09 | -.14 | .18 | .15 |

that form a single dominant factor of Overall Social and Vocational Adjustment (Factor I). One important similarity in findings between these results and the earlier cross-sectional study that does hold up are the two distinct factor patterns of success attributable to enrollees who enter white collar as opposed to blue collar job settings. Thus, the pattern of behaviors for the Overall Social and Vocational Adjustment factor describes a predominantly white collar employee, while Factor II (On-the-Job Success and Satisfaction) applies primarily to those who enter a blue collar work setting. Factor III (Job Search Motivation) provides a relatively coherent picture of the highly motivated ex-enrollee who made every effort to find employment (i.e., substantial positive loadings on number of job sources used, number of places interviewed, number of applications filed, and number of visits to State Employment Service), but tinged with a degree of failure and dissatisfaction (as seen in the negative loadings for job quality, time spent on the job, and the degree to which job expectations were met). The factor bears some resemblance to one that combined a motivational behavior pattern coupled with dissatisfaction found in the cross-sectional criterion development study.

The two remaining criterion factors are apparently specific to behaviors that enter into the job-getting process. One stresses the function of planning (Factor IV Job Planning Competency) and defines the former enrollee who is better able to plan for and seek a better job, take the necessary practical steps in order to do so (e.g., attend school) and, in fact, to have been successful in obtaining a higher level job. The other factor (Short Range Job Orientation, Factor V) depicts the individual who has concentrated on getting a job quickly, holding it, working more hours, and getting a raise but, on the negative side, tending not to be too capable

a planner in terms of his long-term job plans, a willingness to save money, or to attend school. These latter two factors bear little resemblance to any of those identified with the cross-sectional criterion data for a similar post-program sample. However, in any comparisons between these longitudinally obtained post-program factors and those of cross-sectionally derived data, it should be emphasized that gross differences in the dimensions are likely to have resulted from several changes in the criterion variables included or the way in which they could be applied. As examples, the variable of "Number of Salary Raises" which appears in both the present and previous analyses had to be reduced from a continuous variable in cross-sectional use to a dichotomous variable (No/Yes) with the present longitudinal sample because of a highly skewed distribution (i.e., very few enrollees received a salary raise during the comparatively short six-month post-program follow-up period, as contrasted with the better distribution for the variable in the 1 1/2 year post-program span covered by the cross-sectional sample). Similarly "Amount of First Salary Raise" and "Promotions," which appeared as important variables among the previous study criteria, become unusable here (i.e., so few enrollees obtained promotion within six months after leaving the program as to make the variable a meaningless one). G.E.D. as an outcome measure had simply not been included in the cross-sectional analyses.

Although there is some value in being aware of these differences for further efforts at criterion development, the primary interest in this section remains that of the levels of long-term validities found for the ETB measures. Those results are shown in Table 8 in the form of the relationships between each of 22 test scores and the factor scores obtained for the five criterion dimensions. The pattern and magnitude of the test validities, using these

Table 8

Validity Coefficients for Enrollee Test Battery with Post-Program Criterion Factor Scores

Employment Experience (Total Sample: N = 120)

| Measures of Enrollee Test Battery | Criterion Factors I Overall Social & Vocational Adjustment (White Collar) | II On-the-Job Success and Satisfaction (Blue Collar) | III Job Search Motivation | IV Job Planning Competency | V Short-Range Job Orientation |
|---|---|---|---|---|---|
| 1. Job Knowledge | .01 | .15 | .28* | .00 | .16 |
| 2. Self-Esteem | .26* | -.03 | -.07 | .06 | -.11 |
| 3. Def. Gratification | -.02 | -.07 | -.02 | .00 | .04 |
| 4. Level of Aspiration | .04 | -.14 | .03 | .22* | -.24* |
| 5. Level of Plans | .06 | -.21* | -.04 | .08 | -.31* |
| 6. Aspir.-Plans Diff. | -.07 | .02 | .08 | .15 | .01 |
| 7. Att. Toward Authority | .03 | .19* | .06 | .01 | -.04 |
| 8. Job-Seeking Skills | -.05 | .26* | .15 | .05 | .14 |
| 9. Job-Holding Skills | .29* | .08 | -.05 | -.05 | -.04 |
| 10. Motiv. for Voc. Ach. | .12 | .01 | .04 | .05 | -.19* |
| 11. Prac.Reas.(PR): Map Read. | -.17 | .38* | .29* | .06 | .30* |
| 12. PR: Zip Coding | .04 | .28* | .22* | .02 | .16 |
| 13. PR: File Card Sorting[a] | -.05 | .29* | .09 | -.03 | .02 |
| 14. Enrollee Rating Scale | .09 | .12 | -.06 | -.12 | .00 |
| 15. Reading Level(SAT,CAT)[b] | .11 | .02 | -.03 | .12 | .45* |
| Vocational Task Interest | | | | | |
| Aesthetic | .00 | .07 | .04 | .01 | .02 |
| Science | -.26* | .10 | -.10 | .18 | -.17 |
| Technical | -.30* | .27* | -.08 | .01 | -.08 |
| Clerical | .19* | -.08 | -.14 | -.10 | -.01 |
| Business | -.15 | -.04 | -.11 | .01 | .01 |
| Service | .21* | -.07 | .00 | .03 | -.04 |
| Outdoor | -.23* | .26* | .01 | -.05 | .06 |

[a] Males only; N = 58.

[b] N = 53.

*r significant at .05 level.

longer-term outcome dimensions, are obviously improved over the short-term, program completion, validities. In this case, two measures having largely attitudinal components (i.e., little correlation with a variety of cognitive-intellectual measures) and a group of interest scales, achieve a significant level of predictive validity for the dominant criterion dimension, to the virtual exclusion of any validity for cognitive skill measures. Thus, scores on the Self-Esteem and Job-Holding Skills tests are found significantly related to a white collar oriented Overall Social and Vocational Adjustment set of outcomes (r = .26 and .29, respectively). Science, Technical, and Outdoor forms of occupational task interest (interpretable as having a blue collar flavor for the tasks portrayed) are negatively related to that dimension of success (r's = -.26, -.30, and -.23, respectively), while the Clerical and Service interest scales (having a white collar focus) show very low but significant positive relationships (r's = .19 and .21).

By sharp contrast, the validities for the blue collar oriented On-the-Job Success and Satisfaction dimension (Factor II) are consistently highest for precisely those types of tests that have distinct cognitive-reasoning skill requirements, as opposed to the attitudinal ones. The three "Practical Reasoning" measures and the "Job-Seeking Skills" test clearly possess the best validities (r's ranging from .26 for JSS to .38 for PR:Map Reading). Reading Level, surprisingly, does not show any validity for this vocational success dimension, leaving an implication of something more than verbal skill as the major cognitive element involved. Again, by way of contrast, two of the interest scales which were negatively related to the white collar success pattern of Factor I now assume positive validities for the blue collar success pattern characterized by Factor II (i.e., for the Technical Interest scale, r = .27 and for the Outdoor scale, r = .26).

The "Level of Vocational Plans" measure, where it has any degree of significant validity, is seen to be negative in its relationship to later performance. A higher score on "Vocational Plans" is indicative of a poorer post-program score for On-the-Job Success and Satisfaction ($r = -.21$) and the negative relationship holds to an even greater degree for the Short-Range Job Orientation dimension ($r = -.31$). Enrollees whose job intentions are higher (unrealistic?) appear to be those whose levels of job performance and satisfaction are poorer in whatever jobs they do obtain.

A Job Search and Motivation factor (Factor III) is modestly predictable from a triad of ETB measures that are heavily cognitive in their makeup. "Job Knowledge" and two of the "Practical Reasoning" measures (Map Reading and Zip Coding) have the best validities with the highly specific job-seeking efforts that go into defining this dimension. Validity for "PR:Map Reading" is also found to be strong for another specific job performance dimension, that of Short-Range Job Orientation ($r = .30$) and is accompanied by substantial validity for the first, and only, time by the Reading Level score with that same criterion factor ($r = .45$). Interestingly, some additional support is found for the negative aspect of "Vocational Plans" as a job success predictor. That measure joins with "Level of Aspiration" and "Motivation for Vocational Achievement" in their negative validities with Factor V, to indicate that higher levels of job expectations and intentions among a group of disadvantaged adolescents may not be particularly conducive to a narrow range of job-success behaviors (at least in the particular labor market open to these youngsters).

Validities for the post-program criteria, obtainable from a substantial proportion of former enrollees who had never achieved full-time employment

status, must be based on necessarily limited areas of criterion information that are confined to categories of social adjustment, vocational motivation or planning behaviors and personal feelings about employment in general. Lacking specific job-related performance outcomes that can be dealt with, or much in the way of "external" (other-than-enrollee) assessment of capabilities, there are logical limitations in accepting such sharply circumscribed forms of post-program behaviors as representing a suitable range of outcomes. But, despite both logical weaknesses and questionable uses of such a limited criterion domain for these longer-term criteria (including generally poor interpretability and coherence of factors found in previous analyses with the cross-sectional sample), it is felt that a presentation of validity results and a brief discussion may help round out the picture of ETB validities.

In Table 9 the rotated factor loadings are shown for four interpretable factors in the order in which they were extracted from the 29 x 29 matrix of post-program criterion variables that are applicable to those ex-enrollees who do not achieve full-time employment (males and females combined).

Although they are very similar in designation to the group of factors found previously with the cross-sectional data, they are somewhat more coherent and interpretable. Factor I, as the dominant dimension, is readily identified as one of Social Adjustment depicting the ex-enrollee who keeps out of trouble with the police, considers it especially important to do so and also shows a variety of favorable perceptions regarding jobs and job planning. Factor II Success-Oriented Activity combines not only a number of variables that indicate favorable job perceptions but also a higher level of current activity aimed at achieving employment.

Only a few variables define Factor III (High Level of Expectation) which depicts a former enrollee who unquestionably has high salary expectations

Table 9

Rotated Factor Loadings for Post-Program Criteria:
No Employment Experience Group

(Total Group: N = 98)

| Criterion Variables | Factors | | | |
|---|---|---|---|---|
| | I Social Adjustment | II Success Oriented Activity | III High Level of Expectation | IV Work Motivation and Planning |
| 1. # Places Lived Since N.Y.C. | -.02 | -.02 | .30 | .04 |
| 2. Level of Current Activity | .02 | .67 | .16 | .22 |
| 3. Level of Short-Term Plans | -.08 | .72 | .25 | .26 |
| 4. Part-Time Work (No/Yes) | -.09 | .05 | .08 | .55 |
| 5. Job Quality Level Sought | .62 | -.03 | .37 | -.06 |
| 6. Reason Picked Job Sought | .04 | -.02 | .40 | .15 |
| 7. Knowledge of Job Sought | .36 | .26 | -.01 | .19 |
| 8. Ability to Perform Job Sought | .29 | .34 | -.23 | .05 |
| 9. Reason for Ability Judgment | .31 | .36 | .05 | .56 |
| 10. Starting Salary Expected | -.00 | .09 | .83 | .02 |
| 11. Highest Salary Expected | -.04 | .07 | .87 | .02 |
| 12. Ways to Look for Job | .24 | .27 | -.09 | .43 |
| 13. Things to Ask about Job | -.02 | -.10 | -.06 | -.06 |
| 14. Concern for Job Characteristics | .12 | .66 | -.30 | -.31 |
| 15. Long Range Work Plans | .42 | .04 | .20 | .46 |
| 16. Means to Achieve Plans | .08 | -.02 | -.35 | .31 |
| 17. Get Along with Family | -.07 | -.03 | .04 | .38 |
| 18. Future Finan.Contrib.to Family | .05 | -.12 | -.02 | -.48 |
| 19. # People Giving Hard Time | -.44 | .27 | -.16 | -.21 |
| 20. Health Problems | -.12 | .15 | -.03 | -.20 |
| 21. # Times Visited Doctor | -.26 | -.27 | .05 | -.15 |
| 22. Trouble with Police | -.86 | -.03 | .13 | -.02 |
| 23. Imp. of Keep out of Trouble | .47 | .14 | .09 | -.20 |
| 24. Credit Buying | .08 | -.15 | -.01 | .07 |
| 25. Ease of Obtaining Credit | .04 | .33 | .08 | -.18 |
| 26. # Visits to S.E.S. | -.13 | -.05 | -.13 | .03 |
| 27. Obtain G.E.D. Since Leav. Prog. | .01 | -.12 | -.13 | .51 |
| 28. # Actual Police Contacts | -.79 | -.06 | .28 | .03 |
| 29. # Interviewer Contacts Required | -.41 | -.54 | .00 | .01 |

(starting and later salaries) and seeks a higher level job but tends (from the negative loadings on variables #14 and #16) to show evidence of poor planning ability. The final factor, Factor IV, Work Motivation and Planning, is the best representation of the enrollee with ambition, in that he has held part-time employment, obtained the G.E.D., and shows evidence of greater concern with the planning for and seeking of employment.

The validities for the subtests of the ETB, using the factor scores from these four dimensions as the criteria, appear in Table 10. A meaningful pattern of significant validities is found for only one factor and that is the dominant Social Adjustment dimension. Those relationships indicate that tests with heavy stress on cognitive skills possess the highest validity for the social adjustment outcomes in a group that does not obtain full-time post-program employment (r's = .32 for JSS; .20 for PR:Zip Coding; and .38 for PR:File Card Sorting). Several of the vocational interest scales also show levels of significant validity that are consistently negative in their relation to social adjustment (i.e., Aesthetic, r = -.25; Technical, r = -.28; and Outdoor, r = -.23).

Significant validities appearing in Table 10, for the three remaining criterion dimensions, can be seen as scattered and uniformly at such low levels as to provide little benefit for detailed discussion.

C. Employment as Post-Program Criterion

One post-program outcome measure that is often viewed as the best single overriding index of individual or program success is the criterion of full-time employment (No Full-Time Employment vs. Full-Time Employment). That dichotomous outcome variable, however arguable its limitations as a prime form of criterion information, does merit a separate analysis of its role

Table 10

Validity Coefficients for Enrollee Test Battery with Post-Program Criterion Factor Scores

(Former Enrollees; No Employment Experience. Total Group: N = 98)

| Measures of Enrollee Test Battery | Criterion Factors | | | |
|---|---|---|---|---|
| | I Social Adjustment | II Success-Oriented Activity | III High Level of Expectation | IV Work Motivation & Planning |
| 1. Job Knowledge | .14 | .02 | -.13 | -.01 |
| 2. Self-Esteem | .08 | .07 | -.10 | -.06 |
| 3. Deferred Gratification | -.11 | -.19 | .11 | -.05 |
| 4. Level of Aspiration | .09 | -.23* | .16 | -.06 |
| 5. Level of Plans | .17 | -.16 | .08 | -.12 |
| 6. Aspirations-Plans Diff. | -.05 | -.11 | .02 | .21* |
| 7. Attitude Toward Authority | .02 | -.19 | .06 | .16 |
| 8. Job-Seeking Skills | .32* | .12 | .03 | -.01 |
| 9. Job Holding Skills | .14 | .08 | -.13 | -.05 |
| 10. Motivation for Vocational Ach. | -.07 | -.13 | -.06 | -.10 |
| 11. Prac. Reason. (PR): Map Read. | .07 | .01 | .21* | .04 |
| 12. PR: Zip Coding | .20* | .05 | .11 | .01 |
| 13. PR: File Card Sorting[a] | .38* | .17 | .07 | .01 |
| 14. Enrollee Rating Scale | -.10 | .08 | -.01 | -.05 |
| 15. Reading Level (SAT, CAT)[b] | .28 | .28 | .02 | .19 |
| Vocational Task Interest | | | | |
| Aesthetic | -.25* | -.20* | .24* | .12 |
| Science | -.15 | -.14 | .09 | .09 |
| Technical | -.28* | -.03 | .10 | .09 |
| Clerical | .01 | -.15 | -.13 | -.06 |
| Business | -.16 | -.04 | .01 | -.13 |
| Service | -.10 | -.14 | .10 | -.11 |
| Outdoor | -.23* | .04 | .08 | .10 |

[a] Males only; N = 43.

[b] N = 27.

*r significant at .05 level.

for the ETB measures and the extent to which it is a predictable outcome. Accordingly, validity coefficients (as biserial r's) for each of the tests of the ETB, with the employment criterion, are presented in Table 11 for male and female ex-enrollees.

Significant predictive validities, at low to moderate levels for males, are found for three of the tests that are heavily cognitive in their response requirements: Job-Seeking Skills (r = .36), PR:Map Reading (r = .40) and PR:File Card Sorting (r = .24). Validity for the attitudinal scales with the male sample is nil. For females, the significant validities are confined to a single attitudinal scale, "Attitude Toward Authority" (r = .33), and one cognitive skill measure, "PR:Zip Coding" (r = .28). The r of .33, seen for Reading Level, does not quite reach significance with the small sample size available (N = .32), although it is in the appropriate (positive) direction and provides at least a suggestion of potential value for cognitive-verbal skills as a predictor among female enrollees.

This available evidence for the validity of the measures with employment as a criterion serves to bolster the demonstrated effectiveness of several tests that seem to have fairly wide-range validity with other criteria and to guide future choices of ETB measures for desired applications.

Criterion Relevance

The important and very practical intent of identifying the most effective short-term criterion variables must depend, ultimately, on their degree of relationship to more distal (longer-term) measures of performance outcome. To accomplish this, a rectangular matrix of correlations between the 32 program completion (short-term) criterion variables and the 40 post-program

Table 11

Validity Coefficients for Enrollee Test Battery with Post-Program Employment as Criterion

(Employed Full-Time vs. No Full-Time Employment)

| Test | $r_{bis}$ Males (N = 104) | $r_{bis}$ Females (N = 115) |
|---|---|---|
| 1. Job Knowledge | .22 | .01 |
| 2. Self-Esteem | -.03 | .09 |
| 3. Deferred Gratification | .03 | .05 |
| 4. Level of Aspiration | .01 | .02 |
| 5. Level of Plans | -.08 | -.04 |
| 6. Aspirations-Plans Diff. | -.04 | -.05 |
| 7. Attitude Toward Authority | .15 | .33* |
| 8. Job-Seeking Skills | .36* | .15 |
| 9. Job-Holding Skills | .08 | .11 |
| 10. Motivation for Vocational Ach. | -.11 | .18 |
| 11. Practical Reasoning (PR): Map Reading | .40* | .13 |
| 12. PR: Zip Coding | .24* | .28* |
| 13. PR: File Card Sorting | .16 | - |
| 14. Enrollee Rating Scale | -.23 | .09 |
| 15. Reading Level (SAT, CAT)[a] | .02 | .33 |
| Vocational Task Interest | | |
| Aesthetic | -.13 | -.05 |
| Science | .12 | -.15 |
| Technical | -.01 | -.16 |
| Clerical | -.02 | -.02 |
| Business | -.23 | -.15 |
| Service | .18 | .10 |
| Outdoor | .08 | .04 |

[a] N = 47 (Males); N = 32 (Females).

*r significant at .05 level.

criterion variables was drawn up for inspection. Since it is the pattern and level of significant correlations from the resulting 1280 r's that is of basic interest, only significant r's are presented in Tables 12 and 12a. In order to clarify the results further and identify the most important of the short-term criteria having the largest number and magnitude of significant relationships (i.e., the most relevant), these were narrowed down to thirteen such program completion criterion variables displaying eight or more significant relationships.

That "baker's dozen" of the best of the short-term criterion variables is also presented in rank order, based on the number of significant r's (shown in parentheses) for each one, in its relationships to the longer-term post-program outcome variables. Notably it is the short-term variables of highest rank that tend to show relationships of highest magnitude with some of the most important ("objective") of the longer-term outcomes. Ratings by counselors, work site supervisors and peers, keeping out of trouble with police, awareness of job characteristics, confidence in one's vocational future (Vocational Adequacy), and variables of social adjustment (in family and community) show <u>consistent</u> and <u>uniform</u> patterns of significant relationships at substantial levels with such important performance outcomes (six months later) as actual police contacts, ratings of job proficiency by employers, community adjustment, and family adjustment. In every instance these fairly strong correlations are in the expected (logically hypothesized) direction (e.g., the short-term measure of Trouble with the Police is positively related to variables of Actual Police Contacts and negatively to Employer Ratings measured some six months later).

Variables of lesser rank among the 13 most relevant program completion criteria shown (i.e., those having only eight or nine significant correlations

Table 12

Degree of Relevance for 13 "Best" Program Completion Criteria

(Ranked by number of significant r's)*

| Program Completion Criteria | Post-Program Criteria: # Interviewer Contacts Required (1) | # Places Lived Since Left N.Y.C. (2) | Employed Now (No/Yes) (3) | Industry Category (Wh./Bl. Collar) (4) | Job Quality (Level) (5) | Hours Worked per Week (6) | Time on Job (7) | # Jobs Held Since N.Y.C. (8) | Time to Find First Job (9) | # Sources Used for First Job (10) | # Places Interviewed (11) | # Applications Filed (12) | Met Job Expectations (13) | Starting Salary (14) | Salary Raise (No/Yes) (15) | Job Satisfaction (16) | Top Salary Expectation (17) | Level of Short-Term Plans (18) | Job Level Sought (Short Term) (19) | Relevance of Job Choice (20) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Imp. Keep Out of Trouble (19) | | | .22 | -.28 | | | | | .24 | | | | .34 | | | .24 | -.23 | | | .26 |
| Peer Ratings (17) | | | | -.30 | | | .22 | | .24 | | | | | .27 | | .30 | | .20 | | .25 |
| Counselor Rating (17) | | | | | .25 | | .20 | | | | | | | .20 | | | | .21 | .30 | |
| Awareness of Job Characteristics (15) | | | | -.23 | | | | | | -.25 | -.25 | -.20 | | | | .20 | | | | |
| Trouble with Police (14) | | | | .26 | -.22 | | .22 | | | | | | | | .20 | | .41 | -.19 | | |
| Family Feelings (13) | | | | | | | | | | -.22 | -.31 | -.21 | .22 | | | .26 | | | | |
| Feelings of Voc. Adeq. (13) | | | | | .27 | | .26 | | | | -.22 | | .19 | | | | | .19 | .25 | .28 |
| Supervisor Rating (12) | | | | | | | | | | | | | | .22 | | .22 | | | .20 | .24 |
| Imp. Things to Ask Job Interviewer (9) | | | | -.21 | .20 | .25 | | | | | | .22 | | | | | .34 | | .28 | .34 |
| Job Knowledge (9) | .21 | | | | | .22 | | | -.28 | | | | | .37 | | .19 | | | | |
| Take Part-Time Training (9) | -.21 | | | | | | | | .23 | | -.22 | | .23 | | | | | | | |
| Get Along with Family (8) | .24 | -.19 | | | | | | | | | -.22 | | | | | .20 | | | | |
| G. E. D. (8) | -.20 | | | | | | | | | -.23 | -.27 | | .23 | | | | | | | |

*All r's shown are significant at .05 level.

Table 12a

Degree of Relevance for 13 "Best" Program Completion Criteria

(Ranked by number of significant r's)*

| Program Completion Criteria | Post-Program Criteria: Self-Assessment of Job Knowledge (21) | Job Seeking Relevance (22) | Level of Long Range Plans (23) | School Attendance (24) | G. E. D. (25) | Get Along with Family (26) | Family Feelings About Enrollee (27) | Financial Assist. to Family (28) | # People Giving Hard Time (29) | Health Problems (30) | # Times Visited Doctor (31) | Saved Money (32) | Amount Trouble with Police (33) | Imp. Keeping Out of Trouble (34) | Credit Buying (No/Yes) (35) | Ease of Obtaining Credit (36) | # Visits to S.E.S. (37) | Permission for Employer Rating (38) | # Actual Police Contacts (39) | Employer Rating (40) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Imp. Keep Out of Trouble (19) | | .21 | | | | .22 | .42 | .39 | -.39 | | | .24 | -.28 | .42 | .28 | .35 | | | -.43 | .45 |
| Peer Ratings (17) | .23 | | | | | .21 | .31 | | -.33 | | | .29 | -.36 | .22 | | .27 | | | -.25 | .33 |
| Counselor Rating (17) | .22 | .34 | | | .32 | .33 | .38 | | -.35 | | | .20 | -.34 | .22 | .26 | | | | -.33 | .52 |
| Awareness of Job Characteristics (15) | .31 | | | | | | .27 | | -.22 | | | .19 | -.30 | .26 | .24 | | | -.21 | -.28 | .40 |
| Trouble with Police (14) | | | | | | -.26 | | -.25 | .22 | | | | .27 | -.19 | | -.25 | | | .34 | -.32 |
| Family Feelings (13) | | | | .23 | .23 | .23 | .30 | | -.32 | | | | -.33 | | | | | | -.24 | .43 |
| Feelings of Voc. Adeq. (13) | | | .26 | | | | .28 | | -.33 | | | .29 | | | .23 | | | | | .40 |
| Supervisor Rating (12) | | | | | | .27 | .33 | | -.35 | | | | -.27 | | .24 | .26 | | | -.34 | .41 |
| Imp. Things to Ask Job Interviewer (9) | | | | | .19 | | | | | -.24 | | | | | | | | | | |
| Job Knowledge (9) | .21 | | | | -.20 | | | | | | | | | | .21 | | | | | .24 |
| Take Part-Time Training (9) | | | | | .21 | | .21 | | | | | .33 | | | | .27 | | | | .37 |
| Get Along with Family (8) | | | | | | .33 | | | -.32 | | | | -.25 | | | | | | | .35 |
| G. E. D. (8) | | | | .20 | | | | | | | | | | | | .21 | | | -.19 | .35 |

*All r's shown are significant at .05 level.

with longer-term criteria) evidence more scattered and less striking patterns of significant correlation. But even among those five, or so, variables it is worth noting the degree of consistent relationship to such a key criterion as later Employer Rating (r's ranging from .24 for Job Knowledge to .37 for Willingness to Take Part-Time Training). Still other individual correlations or patterns of r's can provide insights into criterion selection or setting of short-term standards as well as clues to subsequent enrollee capability. For example, the three types of judgmental evaluations of enrollee proficiency at program completion (counselor, work site supervisor and peer ratings) are uniformly relevant to later performance criteria of Starting Salary and two family adjustment variables. Similarly, those three short-term ratings, as a set, provide an array of other significant r's for a number of later job-related outcomes dealing with capabilities, satisfactions and planning skills. The many more specific patterns of individual relationships that may be cited in order to answer a variety of specific questions are too numerous to be discussed in detail here. The broader value of the information lies, instead, in the empirical evidence provided for making intelligent (defensible) choices of multiple short-term behavioral outcomes from a group of available measures--and for doing so with flexibility, as a function of data-gathering resources, other research and evaluation constraints. In addition, this best set of the short-term criteria, having a degree of demonstrated relevance, can serve as a starting point for the construction of additional and, perhaps more effective, measures of similar constructs.

Effects of Sample Attrition

Some clues to possible selective biases, resulting from sample attrition over the longitudinal study span, can be sought in the comparisons of test score

Table 13

Comparison of Test Means for Post-Program Sample ("Respondents") and Sample Not Followed Up ("Non-Respondents")

| Test | Respondent Mean (N = 219) | Non-Respondent Mean (N = 344) | Critical Ratio (For Mean Diff.) |
|---|---|---|---|
| 1. Job Knowledge | 19.7 | 19.2 | 1.52 |
| 2. Self-Esteem | 36.2 | 36.6 | 1.21 |
| 3. Deferred Gratification | 50.8 | 52.3 | 2.39* |
| 4. Level of Aspiration | 18.1 | 18.7 | 1.93 |
| 5. Level of Plans | 15.7 | 16.2 | 1.80 |
| 6. Aspirations-Plans Diff. | 3.9 | 3.8 | .42 |
| 7. Attitude Toward Authority | 42.5 | 42.4 | .15 |
| 8. Job-Seeking Skills | 11.1 | 10.8 | .87 |
| 9. Job-Holding Skills | 27.8 | 28.5 | 1.84 |
| 10. Motivation for Vocational Ach. | 38.5 | 39.1 | 1.14 |
| 11. Practical Reasoning (PR): Map Reading | 5.3 | 5.5 | .59 |
| 12. PR: Zip Coding | 7.0 | 7.1 | .27 |
| 13. PR: File Card Sorting[a] | 7.6 | 7.7 | .19 |
| 14. Enrollee Rating Scale | 53.7 | 54.9 | 1.18 |
| 15. Reading Level (SAT, CAT)[b] | 7.7 | 6.5 | 3.38** |
| Vocational Task Interest | | | |
| Aesthetic | 9.8 | 9.9 | .09 |
| Science | 9.3 | 9.7 | 1.85 |
| Technical | 8.7 | 8.9 | .58 |
| Clerical | 10.8 | 11.0 | .74 |
| Business | 10.0 | 10.2 | .71 |
| Service | 10.9 | 11.2 | 1.26 |
| Outdoor | 6.8 | 6.9 | .37 |

[a] Males only; N's = 101 and 151 respectively.

[b] N's = 79 and 81 respectively.

*Significant at .05 level.

**Significant at .01 level.

means shown in Table 13. Critical ratios are presented for the differences between ETB means found for those enrollees who were able to be followed up and interviewed some six months after they left the training program (Respondents), and those who were not later obtainable for inclusion in the post-program sample (Non-Respondents). Only one significant difference ($p < .05$) is found among the 21 ETB test scores and that is for the Deferred Gratification test. Why those enrollees who were followed up should show a significantly lower mean score is difficult to interpret and might best be attributed to the likelihood that one such significant difference would tend to occur at random from among this number of t-tests.

More compelling, as a difference of greater significance between the groups, is the Critical Ratio of 3.38 ($p < .01$) for the Reading Level score, which indicates that those former enrollees who were able to be followed up read at significantly higher grade levels than those who were not able to be included in the post-program sample. Superficial explanations for this particular form of bias can be advanced. For example, those who are more able readers are more easily contacted by the interviewer or perhaps more willing to volunteer to respond to a questionnaire (even though orally presented). Before assigning excessive weight to this potentially biasing influence, however, it should be understood that the finding remains tenuous, since the mean values of the reading score are based on only a minor proportion of the total enrollee sample (N = 160 of the total sample of 563) and may itself provide a biased representation of the originally tested group.

## CONCLUSIONS

From this attempt to identify levels of predictive validity for the Enrollee Test Battery, it is evident--as would be likely for any set of complex behavioral measures--that the validities obtained are largely a function of the particular criteria chosen and the time segment from which those measures of outcome are drawn.

Short-term validation, using criterion dimensions available at the completion of a youth-work training program, resulted in far fewer and less interpretable patterns of validity coefficients than could be achieved with the longer-term, post-program outcomes measurable some six months after the enrollee has left the training program and about one year after the initial testing. What little short-term test validity is found with these samples of disadvantaged, adolescent school-dropouts is differential in its pattern for males and females. Among female enrollees, test validities obtained, using outcomes measured at program completion, appear predominantly for a criterion dimension encompassing variables of overall training program adjustment, whereas, for males, it is a cluster of social adjustment criterion behaviors for which significant validities can best be shown.

Evidence of a more important influence on differential validity than sex could be found, at short-term, in the effects of "employability" (i.e., later ability to obtain full-time employment) as a moderator variable. Those enrollees, who were able to achieve employment after completing the program, represent a subgroup for whom much greater ETB predictive validity could be found using criterion variables at program completion (i.e., they are a more "predictable" subgroup than those who do not obtain later employment). Essentially, two relatively distinct classes of enrollee are implied by such a moderator effect;

a finding which, though of little immediate practical value for test applications, suggests a research need to identify differentiating characteristics of these groups in order to enhance training program methods and the application of individual assessment techniques. Some few of the ETB tests (largely cognitively oriented) showed predictive possibilities in that direction based on their relationships to later employment as an outcome.

It is the more acceptable levels and interpretable patterns of predictive validities found, using longer-term, post-program, performance dimensions, that provide the major source of evidence for the value of the ETB. Coefficients in the .20's and .30's and the overall logic of those relationships serve to support the different approach used in the design of these tests and the value of a number of the behavioral constructs as defensible predictors of a variety of job-related performance outcomes. Although such longer-term validities are logically applicable only to a criterion group who obtain post-program employment, there remains a relatively large sample of former enrollees who fail to obtain full-time employment following their training. For them, the examination of predictive validities that can be attempted must deal with a restricted range of post-program performance outcomes (i.e., nonjob-related adjustments). The limited number of test validities that were found for such a criterion sample apply almost entirely to a "Social Adjustment" dimension of behavioral outcomes and for ETB tests requiring cognitive skill capabilities.

On the basis of the various ETB validation analyses, attempted with this unusually broad spectrum of criterion variables, a number of specific predictor subtests can be defined as acceptable research and evaluation tools for use with youth-work training program populations. By far the most consistent pattern of significant validities is found for those several ETB

measures known to possess a large cognitive skill component. These are the three tests that require the respondent to follow directions in job-related reasoning tasks--i.e., the Practical Reasoning measures of Map Reading, Zip Coding, and File Card Sorting.

As an aid in highlighting the overall validities of each of the three Practical Reasoning measures, concurrent validation results from prior studies (Freeberg, 1970; Freeberg & Reilly, 1972), along with some of the most salient predictive validities found in the present study are summarized in Table 14. The consistent patterns of significant validity found with criteria of counselor and work supervisor proficiency ratings, with a variety of conventional cognitive skill measures and with a number of dominant post-program criterion dimensions, are indicative of a fairly wide range of potential assessment uses for the type of skills tapped by these Practical Reasoning tests.

Surprisingly, the validities of what must be considered a prime measure of cognitive skill--i.e., Reading Grade Level, incorporated in the present study for comparative purposes as a conventional verbal score--had shown itself to be a relatively uneven and unimpressive predictor of many of the important dimensions of performance outcome. But the finding should remain tenuous, since there were severe restrictions in sample sizes and a distinct possibility of any number of unidentified biases in the reading capability of the enrollee subsample for whom the scores were obtainable.

Three measures classified as indicative of enrollee "job orientation" (Job Knowledge, Job Seeking Skills, and Job Holding Skills) are also suitable candidates for a variety of operational uses on the basis of previously obtained concurrent validation data and the levels of post-program predictive validity found in this study. In Table 15, some of the most important

Table 14

Validation Summary for ETB Practical Reasoning (PR) Tests

| Criterion Measures | N | Tests: PR Map Reading | PR Zip Coding | PR File Card Sorting |
|---|---|---|---|---|
| | | Concurrent Validities | | |
| (1) Counselor Ratings | | | | |
| Male | 112 | .24* | .20* | .23* |
| Female | 129 | .06 | .25* | -[b] |
| (2) Work Supervisor Ratings | | | | |
| Male | 102 | .28* | .37* | .31* |
| Female | 128 | .06 | .08 | - |
| (3) Paragraph Comprehension (Stanford Achievement Test) | 70 | .52* | .61* | - |
| (4) Vocabulary (ABLE) | 70 | .42* | .41* | - |
| (5) Arithmetic (ABLE) | 70 | .45* | .47* | - |
| (6) Figure Analogies | 70 | .50* | .54* | - |
| (7) Choose-A-Path (Maze Test) | 70 | .38* | .29* | - |
| | | Predictive Validities | | |
| (1) Counselor Ratings[a] (Program Completion) | 155 | .11 | .19* | .21* |
| (2) Work Supervisor Ratings[a] (Program Completion) | 155 | .11 | .18* | .15 |
| (3) Job Success and Satisfaction Dimension (Post-Program) | 120 | .38* | .28* | .29* |
| (4) Job Search Motivation Dimension (Post-Program) | 120 | .29* | .22* | .09 |
| (5) Employed vs. Not Employed (Post-Program) | | | | |
| Male | 104 | .40* | .24* | .16 |
| Female | 115 | .13 | .28* | - |

[a] Males only.

[b] Not used.

*Significant at .05 level or better.

Table 15

Validation Summary for ETB Job-Orientation Measures

| Criterion Measures | N | Tests: Job Knowledge | Tests: Job Holding Skills | Tests: Job Seeking Skills |
|---|---|---|---|---|
| | | **Concurrent Validities** | | |
| (1) Counselor Ratings | | | | |
| Male | 112 | .22* | .22* | .32* |
| Female | 129 | .19* | .26* | .26* |
| (2) Work Supervisor Ratings | | | | |
| Male | 102 | .24* | .19 | .35* |
| Female | 128 | .03 | .17 | .22* |
| (3) Paragraph Comprehension (Stanford Achievement Test) | 70 | .44* | .34* | .75* |
| (4) Vocabulary (ABLE) | 70 | .32* | .17 | .57* |
| (5) Arithmetic (ABLE) | 70 | .37* | .44* | .58* |
| (6) Figure Analogies | 70 | .37* | .12 | .42* |
| (7) Choose-A-Path (Maze Test) | 70 | .29* | .12 | .32* |
| | | **Predictive Validities** | | |
| (1) Counselor Ratings[a] (Program Completion) | 155 | .10 | .20* | .23* |
| (2) Work Supervisor Ratings[a] (Program Completion) | 155 | .16* | .24* | .17* |
| (3) Job Success and Satisfaction Dimension (Post-Program) | 120 | .15 | .08 | .26* |
| (4) Job Search Motivation Dimension (Post-Program) | 120 | .28* | -.05 | .15 |
| (5) Overall Social & Vocational Adjustment Dimension (Post-Program) | 120 | .01 | .29* | -.05 |
| (6) Employed vs. Not Employed (Program Completion) | | | | |
| Male | 104 | .22 | .08 | .36* |
| Female | 115 | .01 | .11 | .15 |

[a]Males only.

*Significant at .05 level or better.

concurrent and predictive validation results are featured in an overview of these three measures.

Taken together, the three measures of job orientation, specific to knowing what the vocational world is about and how to function properly within it, have displayed validities over a number of important areas of behavioral outcome. Clearly the Job Seeking Skills Test, with its largely cognitive demands, has the widest range validity and rivals the best of the Practical Reasoning measures in value. The Job Holding Skills Test (with its focus on understanding appropriate on-the-job behaviors) possesses a stronger attitudinal than intellectual component and remains a measure that can also be said to hold its own as a distinctly worthwhile assessment tool.

Of the purely attitudinal measures, only two would appear to possess sufficient value to warrant highlighting available concurrent and predictive validities. These are the measures of Self-Esteem and Attitude Toward Authority which, from Table 16, can be seen to possess generally spotty patterns of validity. As would be expected of any attitudinal measures, the two have been found to bear only nominal relationships to the cognitive intellectual tests used as concurrent criteria (i.e., they are comparatively independent constructs). Interestingly, however, where some overlap does exist, it occurs exclusively for the ability to discern "appropriate" (i.e., middle-class) behaviors toward authority figures. Certainly, the predictive value of these scales is far more limited than the six measures previously summarized. But each can be considered distinct candidates for retention and limited application--at least until additional validation information with better quality criterion data can be made available.

Table 16

Validation Summary for ETB Attitudinal Measures

| Criterion Measures | N | Tests: Self-Esteem | Tests: Attitude Toward Authority |
|---|---|---|---|
| **Concurrent Validities** | | | |
| (1) Counselor Ratings | | | |
| Male | 112 | .04 | .12 |
| Female | 129 | .15 | .06 |
| (2) Work Supervisor Ratings | | | |
| Male | 102 | .01 | .22* |
| Female | 128 | .21* | .06 |
| (3) Paragraph Comprehension (Stanford Achievement Test) | 70 | -.16 | .31* |
| (4) Vocabulary (ABLE) | 70 | -.05 | .12 |
| (5) Arithmetic (ABLE) | 70 | -.03 | .27* |
| (6) Figure Analogies | 70 | -.19 | .19 |
| (7) Choose-A-Path (Maze Test) | 70 | .02 | .28* |
| **Predictive Validities** | | | |
| (1) Counselor Ratings[a] (Program Completion) | 155 | .20* | .06 |
| (2) Work Supervisor Ratings[a] (Program Completion) | 155 | .13 | .08 |
| (3) Overall Social & Vocational Adjustment Dimension (Post-Program) | 120 | .26* | .03 |
| (4) Job Success and Satisfaction Dimension (Post-Program) | 120 | -.03 | .19* |
| (5) Employed vs. Not Employed (Post-Program) | | | |
| Male | 104 | -.03 | .15 |
| Female | 115 | .04 | .33* |

[a]Males only.

*Significant at .05 level or better.

It is evident that the remaining attitudinal tests (Motivation for Vocational Achievement, Deferred Gratification, the Enrollee Rating Scale, and the score based on Aspirations minus Plans), given their present content and/or scoring methods, have shown little justification on the basis of their predictive validation results to justify continued application with a youth-work training program population.[1]

As a group the Vocational Interest scales would appear to be worth retaining. Not only has their pattern of predictive relationships proven revealing--even for the very broad-based outcomes used here--but other evidence, in a study by Lokan (1973), suggests that rescoring of the scales (based on factor analyses with samples of disadvantaged vocational high school students) can result in acceptable validities with criteria of success in different training curricula. At minimum, there is a justification for attempting more "precise" validation of the interest scales for work training program enrollees, using criteria of actual entry into occupations, as well as achievement within different occupational groups (impractical alternatives with the present study sample and available data).

The measures dealing with Level of Plans and Aspirations had been found to possess only scattered significant predictive validities that are at times difficult to interpret and suggestive of considerable caution in any attempted uses pending further validation evidence.

It is a truism that the validity coefficients obtained and the conclusions regarding the value of any ETB measures can be no better than the

---

[1] Available time and resources, however, did not permit attempts to assess several different scoring techniques that might have improved the validities for these measures.

quality of the criterion data obtained. The point is worth amplifying here. Marked longitudinal sample attrition, to an extent where the resulting sample sizes made analyses by sex infeasible for long-term validation, widespread instances of missing information--often for key criterion variables, and the possibility of selective biases in a hard-to-locate sample of poverty-level adolescents (e.g., a finding of higher levels of reading skill in an enrollee group that could be followed up) all represent undesirable features that may have distorted the validation results to some extent. Despite the flaws, however, it is felt that the seven to nine subtests that can be said to have "survived," by virtue of their display of significant validities over a variety of outcome dimensions, could justifiably be applied to any number of research and evaluation needs for youth-work training programs.

Further attempts at improving validity for the best of the ETB measures are certainly in order. The success with which the intent can be achieved is dependent not only on improving sample quality and information availability, but also by obtaining better knowledge of the value (relevance) of alternative criterion measures. In turn, that knowledge permits a sharpening of criterion choices for test validation. As seen in the present study, identification of a relevant set of program-completion criterion measures, based on their relationships to later performance outcomes, serves to depict a meaningful, quantitative translation of program objectives over the short term. In a broader sense, such information also helps to define program objectives worth measuring, at any level, based on their predictive relevance and their predictability [e.g., indices of health

status used in the present study appear neither especially predictive (relevant) as short-term outcomes, nor predictable as longer-term objectives].

This initial attempt to determine test validities and criterion relevance in a predictive framework provides a flexible core of assessment tools of demonstrable value, as a measurement structure to be used or built on. Additional efforts to develop and apply more relevant criteria, in order to enhance test validity, constitute a continuing and interrelated process that should be seen as feeding all aspects of research and evaluation need. Only through that process can defensible measures be made available to permit better understanding of program design, operation, and effectiveness.

Implications for Future Use of the Measures

Any set of behavioral predictors, or tests, prove (and improve) their value only by continuous application in a variety of circumstances that allow for use of converging criterion constructs. Consequently, further applications of the measures of the Enrollee Test Battery should not only serve as a means to satisfy the immediate assessment and information needs of youth-work training programs but should combine, along with such use, the collection of appropriate research data that add to needed developmental improvements in the measures themselves.

Two obvious avenues are open for practical use and further development with the existing measures. The first can be designated as an evaluative mode which entails use of the ETB to obtain information on a group or normative basis. This is probably the most justifiable way in which to apply the seven to nine "surviving" ETB Tests based on their currently demonstrated measurement properties. In general, modest levels of validity and even marginal

reliabilities (e.g., $r_{tt's}$ in the .50's) can be considered suitable for the fairly broad trends and normative conclusions sought from evaluative and research data obtained in a youth-work training setting.

Application of the ETB measures in this mode would be concerned primarily with the detection of intraprogram effects on enrollee behavioral characteristics resulting from training program participation (e.g., with some type of "before and after" study design to measure treatment effects during the course of the program). Information at this level is likely to be desired by project directors and such professionals as counselors, work supervisors, skills instructors, etc., in order to make relatively rapid decisions about the effectiveness of different instructional methods and guidance techniques (and is also most likely to be within the scope of their available evaluation resources). For such a purpose, the measures of the ETB that most readily meet acceptable standards are (a) the Practical Reasoning Tests (i.e., the Map Reading, Zip Coding, and File Card Sorting that assess direction-following skills in simulated job tasks); (b) the vocational orientation measures of Job Knowledge, Job Holding Skills, and Job Seeking Skills (dealing with job demands and the means of finding and retaining employment); and (c) the attitudinal measures of Self-Esteem and Attitude Toward Authority (dealing with perceptions of self-worth and appropriateness of various social-interpersonal behaviors). The suitability of these eight measures for any evaluative uses lies not only in their demonstrated degree of measurement value but in their content as well, since they represent a set of behavioral constructs and skills amenable to modification by reasonable training procedures.

The program completion and post-program criterion measures identified in the present study also merit (and benefit from) further application--along

with the ETB measures--in an evaluative framework. The most immediate benefit realizable from the criteria, as "end-point" information, lies in their use for making comparative estimates of effectiveness between youth-work training projects or programs that possess clearly definable differences in their training approach (model) or their curriculum components (e.g., amount of remedial skills training, amount of time spent in counseling, etc.). Thus, a number of youth-work training projects, using a similar type of training model and target population, may be contrasted on the relatively short-term behavioral outcomes that define "success" at program completion (probably a high priority form of evaluative information for an administrator, such as a Regional Manpower Administrator, who must identify the most effective projects under his jurisdiction). The dominant dimensions, or clusters of successful behaviors at program completion that contained the most relevant criterion variables for this level of information need, were the ones that dealt with overall adjustment in the training program, vocational and social attitudes, community and interpersonal adjustments, and levels of willingness to plan for and seek employment.

Similarly, at still another (higher) level of information need (e.g., governmental policy makers with larger scale data-gathering resources) different types of youth-work training programs, with somewhat different formats and training content, could be contrasted on the basis of broader criterion dimensions involving longer-term, post-program patterns of behavioral outcomes. Such criteria were found in clusters of independent post-program dimensions that define demonstrated "success" of former enrollees in terms of social adjustments, specific aspects of job performance, skills in planning to find jobs, and their degree of motivation to search for employment.

Any examinations of differences or contrasts between training projects and programs, as suggested above, are sensible only when the evaluation design allows study samples to be matched in pertinent ways on initial status. Ideally, where a number of ETB measures have been applied in "before and after" comparisons, the same enrollee samples can continue to be followed up longitudinally for later contrasting evaluations of program effectiveness. Those ETB measures then serve as the needed covariates for control on initial enrollee status. By incorporating the predictor and criterion measures in some such comprehensive data-gathering strategy, the results are not only usable: (a) to define intratraining project changes for enrollee characteristics over the course of training, and (b) to contrast interproject or interprogram differences on short-term and longer-term outcomes, they also provide an inescapable and important data by-product of developmental benefit to the ETB measures. That by-product is the repetition of the validation process (i.e., cross-validation), which allows for confirmation of the predictive validities obtained in the initial findings of the present study.

The second major approach to use of the ETB measures would be that of a guidance mode in which information from test scores are applied in order to make decisions for (or by) the individual enrollee (e.g., assignment to a specific curriculum; emphases on certain areas of guidance need).

Although any measure can serve multiple functions, there are precautions applicable to guidance uses of tests which dictate generally more stringent demands on their measurement properties. Thus, not all of the eight ETB measures that were cited as appropriate to a group evaluative function would, as yet, meet conventional standards for levels of reliability and validity

that justify unrestrained use in making guidance decisions. Three of the measures in the battery that can be defended, however, as strong candidates for guidance use with work-training program enrollees, are the two Practical Reasoning Tests of Map Reading and Zip Coding, and the Job Seeking Skills Test. The three are distinctly cognitive in content and suggestive of specific areas of guidance or instructional need. To a somewhat lesser extent, the Job Holding Skills Test with its predominantly attitudinal focus (and suggestive of still other behaviors for remediation) also displays measurement qualities sufficient to justify its use for guidance purposes.

Given information regarding percentile norms by sex, recommended cutoff scores derived from available validities and directions for administration and scoring (i.e., conventional test manual information), these four tests of the ETB would have as much, if not greater, practical value--based on their demonstrated predictive validity for a wide range of criteria--than any presently available published measures which claim to be applicable to educationally disadvantaged adolescents or young adults.

Use of the measures in a guidance mode can also be coupled with further developmental improvements. For example, the overall value of the tests as aids in making guidance or placement decisions might be determined on the basis of enrollee "success", following training curriculum or program assignments, with and without the availability of information from the ETB measures. Data available from such a study could, at the same time (given samples of sufficient size with corrections for range restriction), allow for examination of specific differential predictive validities of each of the tests for different curricula, training techniques, or training programs.

Whether through normative (group) applications, for programmatic evaluative purposes, or as aids in providing information for guiding the individual

enrollee, the ETB and the criteria constitute a defensible measurement foundation from which operational benefits can proceed to be realized in the form of corrective information fed back to programs and in improved research knowledge of the enrollee population served.

## References

Borus, M. E., & Tash, W. R. Measuring the impact of manpower programs: A primer. Ann Arbor, Mich.: Institute of Labor and Industrial Relations, 1970.

Freeberg, N. E. Development of evaluation measures for use with neighborhood youth corps enrollees. Project Report 68-5. Princeton, N. J.: Educational Testing Service, 1968.

Freeberg, N. E. Assessment of disadvantaged adolescents: A different approach to research and evaluation measures. Journal of Educational Psychology, 1970, 61, 229-240.

Freeberg, N. E., & Reilly, R. R. Development of guidance measures for youth-work training program enrollees. Phase I: Measurement of program objectives and the development of criteria. Project Report 71-13. Princeton, N. J.: Educational Testing Service, 1971.

Freeberg, N., & Reilly, R. Validation of a test battery for youth work training program enrollees. Research Memorandum 72-11. Princeton, N. J.: Educational Testing Service, 1972.

Ghiselli, E. E. Moderating effects and differential reliability and validity. Journal of Applied Psychology, 1963, 47, 81-86.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Kelley, T., Madden, R., Gardner, E., & Rudman, H. Stanford Achievement Test: Form W, Intermediate I. New York: Harcourt, Brace and World, 1964.

Levitan, S. A., & Mangum, G. L. Federal training and work programs in the sixties, Part V. Ann Arbor, Mich.: Institute of Labor and Industrial Relations, 1969.

Lokan, J. The development and preliminary examination of an instrument to measure the vocational interests of low-reading high school students. Paper presented at N.C.M.E. Annual Conference, February 1973.

Thorndike, R. L. Personnel selection. New York: Wiley, 1949.

Tiegs, E., & Clark, W. California Achievement Tests. Level 4. New York: California Test Bureau/McGraw Hill, 1970.

United States Department of Labor. A compilation of papers on data gathering in manpower research. (Unpublished manuscript.) Washington, D.C.: Office of Manpower Research, September 1969.

Wherry, R. J. The past and future of criterion evaluation. Personnel Psychology, 1957, 10, 1-5.

APPENDIX A

PROGRAM COMPLETION CRITERION VARIABLES

Criterion Variables

1. G.E.D.: Working toward or received general equivalency diploma (No/Yes)

2. Voc. Ad.: Feelings of vocational adequacy scale (Total score on 7-item scale)

3. 1st Move: First job intention after leaving N.Y.C. (Level of response rated on 4-point scale)

4. Reason 1st Move: Rating of relevance of response

5. Job Awareness: Degree of concern with conventional aspects of jobs such as salary, raises, security (total score on 7-item scale)

6. Job Knowledge: Enrollee estimate of knowledge of requirements for preferred job

7. Ability to Perform Job: Enrollee estimate of ability to perform preferred job, single item, 3-point scale

8. Starting Salary Expected: (in preferred job)

9. Top Salary Expected: (in preferred job)

10. Ways to Look for Job: # Relevant responses

11. Important Things to Ask Job Interviewer: # Relevant responses

12. Short-Term Job Plans: Degree of suitability, single item, 3-point scale

13. Long-Term Job Plans: Rated quality level of job anticipated 5 to 10 years later

14. Ability to Accomplish Plans: # Relevant responses

15. Get Along with Family: Enrollee estimate, single item, 3-point scale

16. Family Feelings About Enrollee: Enrollee estimate of family feeling about how enrollee is doing, single item, 3-point scale

17. Financial Assistance to Family: Amount given

18. Save Money: Single item, 3-point scale

19. # People Giving Hard Time: Number of categories of people checked by enrollee from 13 categories (e.g., work supervisor, storekeepers, police, etc.)

20. Health Problems: (No/Yes)

21. Trouble with Police: Enrollee estimate of amount of trouble since enrollment in N.Y.C.

22. Importance of Keeping out of Trouble: Single item, 3-point scale

23. Number of Jobs Chosen for Training: Number chosen by enrollee of jobs in which he is willing to accept training

24. Willingness to Train Full Time: Amount of required pay indicated

25. Willingness to Train Part Time: Amount of required pay indicated

26. Work Motivation: Total score on 7-item scale

*27. # Work Sites: # Site assignments during N.Y.C. enrollment

*28. # Days Absent: From training worksite(s)

*29. Counselor Rating: Total score on 11-item scale

*30. Work Supervisor Rating: Total score on 10-item scale

*31. Peer Rating: Mean rating score on 6-item scale for peer known socially and within training program. Mean score for 4-item scale for peer known "outside of N.Y.C. job only"

*32. # Times Chosen for Peer Rating

---

*Data from sources other than enrollee response.

APPENDIX B

POST-PROGRAM CRITERION VARIABLES

(Employment Experience Group)

Criterion Variables

*1. # Interviewer Contacts Required: (To find enrollee respondent)

2. # Places Lived Since Left N.Y.C.

3. Employed Now: (No/Yes)

4. Industrial Category: (Wh. Collar/Bl. Collar)

5. Job Quality (Present or Lost job): Level based on job title (3-point scale Lo to Hi)

6. Hours Worked/Wk.

7. Time on Present (or Last) Job

8. # Jobs Held Since N.Y.C.

9. Time to Find 1st Job

10. # Sources Used to Find 1st Job

11. # Places Interviewed

12. # Applications Filed

13. Met Job Expectations: Single item, 3-point scale

14. Starting Salary

15. Salary Raise: (No/Yes)

16. Job Satisfaction: Total score on 9-item scale

17. Top Salary Expect.: (Present job)

18. Level of Short-Term Plans: Rated on 4-point scale

19. Job Level Sought (Short-Term): Rated on 4-point scale

20. Reason for Job Selection: (Irrel./Relevant)

21. Knowledge of Job Sought: Respondent estimate, single item, 3-point scale

22. Steps to Obtain Desired Job: # Relevant responses

23. Level of Long-Range Work Plans
24. Amt. School Attendance: Single item, 3-point scale
25. Obtained G.E.D. Since Leaving N.Y.C.: (No/Yes)
26. Getting Along with Family: Single item, 3-point scale
27. Family Feelings About Enrollee: Single item, 3-point scale
28. Financial Assistance to Family
29. # People Giving Respondent Hard Time
30. Health Problems: (No/Yes)
31. # Times Visited Doctor
32. Save Money: 3-point scale from lo to hi
33. Amount Trouble with Police: Respondent estimate since leaving N.Y.C.
34. Importance of Keeping out of Trouble: Single item, 3-point scale
35. Credit Buying: (No/Yes)
36. Ease of Obtaining Credit: Single item, 3-point scale
37. # Visits to U. S. Employ. Service
38. Permission for Employer Rating: (No/Yes)
*39. # Actual Police Contacts: Since leaving N.Y.C.
*40. Employer Rating: 4-item scale

POST-PROGRAM CRITERION VARIABLES

(No Employment Experience)

Criterion Variables

*1. # Places Lived Since Left N.Y.C.
2. Level of Current Activity: (Rated on 4-point scale dealing with job searching and school)
3. Level of Short-Term Plans: (Rated on 4-point scale)

---

*Data from sources other than enrollee response.

4. Part-Time Work: (No/Yes)
5. Job Quality Level Sought: (Rated on 4-point scale)
6. Reason Picked Job Sought: (Irrel./Rel.)
7. Knowledge of Job Sought: (Respondent estimate) Single item, 3-point scale
8. Ability to Perform Job Sought: Single item, 3-point scale
9. Reason for Ability Judgment: (# Relevant responses)
10. Starting Salary Expected
11. Highest Salary Expected
12. Ways to Look for Job: (# Relevant responses)
13. Things to Ask About Job: (# Relevant responses)
14. Concern for Job Characteristics: Total score on 7-item scale
15. Long-Range Work Plans: (Quality level rating)
16. Means to Achieve Plans: (# Relevant responses)
17. Get Along with Family: Single item, 3-point scale
18. Future Financial Contrib. to Family
19. # People Giving Hard Time
20. Health Problems: (No/Yes)
21. # Times Visited a Doctor
22. Trouble with Police: (Respondent estimate)
23. Importance of Keeping out of Trouble
24. Credit Buying: (No/Yes)
25. Ease of Obtaining Credit: Single item, 3-point scale
26. # Visits to State Employment Service
27. Obtain G.E.D. Since Leaving Program: (No/Yes)

*28. # Actual Police Contacts

*29. # Interviewer Contacts Required

---

*Data from sources other than enrollee response.

APPENDIX C

## FORMULA FOR STANDARDIZED FACTOR SCORES

(When there is missing data)

$$\underset{Nxm}{F} = \underset{(Nxp)}{Z} \quad \underset{(pxm)}{S} \quad \underset{(mxm)}{(S'S)^{-1}}$$

$$\underset{Nxm}{ZS} = [f_1 f_2 \ldots f_m]$$

$$f_{ij} = p\left(\frac{\sum_{i=1}^{p} \delta_i \, z_i \, a_i}{\sum_{i=1}^{p} \delta_i}\right)$$

that is $f_{ij}$ is the first factor score for the jth person and:

Z is an Nxp matrix of Z scores

S is a pxm rotated factor matrix

$\delta_i$ is equal to 1 when $z_i$ is present and 0 otherwise